



(72) FOURNIER, ALAIN, FR

(72) MERCKEN, LUC, FR

(72) PRADIER, LAURENT, FR

(72) PAUL, MARIE-FRANCOISE, FR

(71) AVENTIS PHARMA S.A., FR

(51) Int.Cl.⁶ C12N 15/12, A61K 31/70, C12Q 1/68, A61K 38/17, G01N 33/68,
C07K 14/47, A61K 39/395, C07K 16/18

(30) 1997/10/01 (97 12221) FR

(30) 1998/07/31 (60/094.933) US

(54) **ACIDES NUCLEIQUES CODANT POUR DES PROTEINES
CAPABLES D'INTERAGIR AVEC LES PRESENILINES**

(54) **NUCLEIC ACIDS CODING FOR PROTEINS CAPABLE OF
INTERACTING WITH PRESENILINS**

(57) La présente invention concerne de nouvelles séquences nucléotidiques et peptidiques, et leur utilisation pharmaceutique. Plus particulièrement, la présente invention concerne des acides nucléiques codant pour des protéines capables d'interagir avec les présénilines.

(57) The invention concerns novel nucleotide and peptide sequences, and their pharmaceutical use. More particularly, it concerns nucleic acids coding for proteins capable of interacting with presenilins.

PCTORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/12, C07K 14/47, 16/18, A61K 38/17, C12Q 1/68, A61K 39/395, 31/70, G01N 33/68	A1	(11) Numéro de publication internationale: WO 99/16874 (43) Date de publication internationale: 8 avril 1999 (08.04.99)
(21) Numéro de la demande internationale: PCT/FR98/02093 (22) Date de dépôt international: 30 septembre 1998 (30.09.98) (30) Données relatives à la priorité: 97/12221 1er octobre 1997 (01.10.97) FR 60/094,933 31 juillet 1998 (31.07.98) US (71) Déposant (pour tous les Etats désignés sauf US): RHONE-POULENC RORER S.A. [FR/FR]; Direction Brevets, 20, avenue Raymond Aron, F-92160 Antony (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): FOURNIER, Alain [FR/FR]; 28, avenue Roger Salengro, F-92000 Chatenay Malabry (FR). MERCKEN, Luc [FR/FR]; 18 bis, rue Jules Joffrin, F-94100 Saint Maur (FR). PRADIER, Laurent [FR/FR]; 5 bis, rue d'Alésia, F-75014 Paris (FR). PAUL, Marie-Françoise [FR/FR]; 5, rue de l'Insurrection Parisienne, F-94600 Choisy-le-Roi (FR). (74) Mandataire: LANCELOT, Géraldine; Rhône-Poulenc Rorer S.A., Direction Brevets, 20, avenue Raymond Aron, F-92165 Antony Cedex (FR).		(81) Etats désignés: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HU, ID, IL, IS, JP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, brevet ARIPO (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Publiée <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont requises.</i>
(54) Title: NUCLEIC ACIDS CODING FOR PROTEINS CAPABLE OF INTERACTING WITH PRESENILINS (54) Titre: ACIDES NUCLEIQUES CODANT POUR DES PROTEINES CAPABLES D'INTERAGIR AVEC LES PRESENILINES (57) Abstract The invention concerns novel nucleotide and peptide sequences, and their pharmaceutical use. More particularly, it concerns nucleic acids coding for proteins capable of interacting with presenilins. (57) Abrégé La présente invention concerne de nouvelles séquences nucléotidiques et peptidiques, et leur utilisation pharmaceutique. Plus particulièrement, la présente invention concerne des acides nucléiques codant pour des protéines capables d'interagir avec les présénilines.		

ACIDES NUCLEIQUES CODANT POUR DES PROTEINES CAPABLES D'INTERAGIR AVEC LES PRESENILINES

La présente invention concerne de nouvelles séquences nucléotidiques et peptidiques,
5 et leur utilisation pharmaceutique. Plus particulièrement, la présente invention
concerne des acides nucléiques codant pour des protéines capables d'interagir avec les
présénilines, les procédés de production des dites protéines ainsi que des
compositions les contenant.

La maladie d'Alzheimer est la plus courante des démences. Sa fréquence
10 globale augmente avec l'âge, elle est de 5 à 10% après 65 ans et de 20% au delà de
80 ans.

Actuellement, il n'existe aucun traitement spécifique et seuls des soins
palliatifs peuvent être offerts aux patients. Cette maladie n'est vraiment confirmée
qu'après l'analyse histologique du cerveau en *post-mortem* par manque d'un
15 diagnostic précis et précoce.

La maladie d'Alzheimer est liée à des lésions cérébrales organiques
constituant des signes histologiques: plaques séniles, dégénérescences
neurofibrillaires neuronales, perte et atrophie des neurones au niveau du cortex.

L'étiologie de la maladie est complexe, de nombreux gènes y sont impliqués
20 et les différents mécanismes physiopathologiques auxquels ils participent ne sont pas
entièrement compris.

Il a été décrit des altérations anatomiques du cerveau observées chez des
patients de moins de 65 ans, liées aux démences préséniles. Actuellement, on englobe
sous le terme "démences de type Alzheimer", les démences préséniles (avant 65 ans)
25 aussi bien que les démences séniles (après 65 ans) présentant les mêmes
caractéristiques.

Il existe deux formes de la maladie : des formes sporadiques, c'est-à-dire non
liées à l'hérédité et des formes familiales.

Les formes sporadiques de la maladie sont les plus fréquentes, elles
30 concernent 60 à 90% des patients et apparaissent le plus souvent après 65 ans. Les

origines de la pathologie demeurent inconnues dans la plupart des cas. Cependant, il a été établi que des traumatismes crâniens, des inflammations, des chocs affectifs ou des situations de grand stress augmentent la probabilité d'apparition de la maladie (Seubert *et al.*, 1992).

5 Les formes familiales de la maladie représentent 10 à 40% des cas. Elle sont classées en deux groupes, les formes tardives (après 65 ans) et les formes précoces (avant 65 ans). Dans les cas précoces, l'affection se transmet selon un caractère dominant et autosomal.

10 En 1992, l'existence d'un locus AD3 a été mise en évidence sur le chromosome 14 et en 1995 le gène de la préséniline 1 (PS-1) a été isolé (le gène a été initialement appelé S182) (Sherrington *et al.*, 1995). Des mutations du gène de la préséniline 1 sont impliquées dans 70% des formes très précoces de la maladie (avant 55 ans). De très rare cas de formes précoces sont liées à des mutations du gène de la préséniline 2 situé sur le chromosome 1 (Levy-Lahad *et al.* 1995). Ce gène présente
15 67% d'identité avec le gène de la préséniline 1.

Le gène est formé de 12 exons, les exons 3 à 12 codent pour la protéine. Il est exprimé de façon ubiquitaire et génère un transcript majoritaire de 2,7 kb et un transcript minoritaire de 7,5 kb (Kovacs, 1996). La protéine, comprend 467 acides aminés, et présente la structure d'une protéine membranaire. Elle présente au moins 8
20 domaines transmembranaires potentiels. Les acides aminés correspondant aux codons touchés par les mutations répertoriées se répartissent sur l'ensemble de la protéine, ils se situent aussi bien dans les régions hydrophobes (TM) que dans les boucles hydrophiles (BL), avec cependant, deux régions privilégiées, TM2 et BL6.

Actuellement, plus de 35 mutations ponctuelles ont été rapportées et les
25 mutations identifiées jusqu'à présent sont presque toutes de type "faux sens" (c'est à dire qu'elles aboutissent au changement ponctuel d'un acide aminé sur la séquence de la protéine). Toutefois, un autre type de mutation affectant l'épissage a récemment

été rapporté (Perez-Tûr *et al.*, 1995), dans ce cas, le gène est altéré par la perte de l'exon 9.

Le rôle des présénilines dans la cellule normale est encore mal connu. De nombreuses hypothèses ont été émises. Parmi ces hypothèses, la préséniline 1 (PS-1) pourrait avoir un rôle dans le trafic cellulaire. Cette hypothèse est confortée par un résultat immunocytochimique qui a permis de localiser PS-1 au niveau du réticulum endoplasmique et de l'appareil de Golgi dans des cellules de neurogliome humain (Kovacs *et al.*, 1996). Par ailleurs, PS-1 jouerait un rôle dans la communication cellulaire et les mécanismes d'endocytose et la protéolyse de l'APP (Dewji et Singer, 1996). Une troisième hypothèse implique PS-1 dans des interactions avec le cytosquelette et notamment les microtubules associés à la protéine Tau (Lew et Wang, 1996). Enfin, compte tenu de sa structure, PS-1 pourrait jouer le rôle d'un récepteur couplé aux protéines G participant à la transduction du signal ou encore celui d'un transporteur ou d'un canal ionique (Van Broeckhoven, 1995).

La recherche de partenaires interagissant avec les présénilines permettrait de comprendre le rôle exacte des présénilines.

La présente invention résulte de la mise en évidence par la demanderesse d'acides nucléiques codant pour des protéines capables d'interagir avec les présénilines. La mise en évidence de tels acides nucléiques permet d'envisager la préparation de nouveaux polypeptides utilisables pharmaceutiquement.

Un premier objet de l'invention concerne donc des séquences nucléotidiques codant pour des protéines capables d'interagir avec les présénilines.

Selon un mode particulier, les séquences nucléotidiques selon l'invention codent pour des protéines ou des polypeptides capables d'interagir avec tout ou partie de la grande boucle de la préséniline PS-1.

Plus préférentiellement les séquences nucléotidiques selon l'invention comprennent tout ou partie de la SEQ ID N°1 ou de ses dérivées ou tout ou partie de la SEQ ID N°2 ou de ses dérivées.

Au sens de la présente invention, le terme séquence nucléotidique dérivée
5 désigne toute séquence différant de la séquence considérée en raison de la
dégérescence du code génétique, obtenue par une ou plusieurs modifications de
nature génétique et/ou chimique, ainsi que toute séquence hybridant avec ces
séquences ou des fragments de celles-ci et codant pour un polypeptide selon
l'invention. Par modification de nature génétique et/ou chimique, on peut entendre
10 toute mutation, substitution, délétion, addition et/ou modification d'un ou plusieurs
résidus. Le terme dérivé comprend également les séquences homologues à la
séquence considérée, issues d'autres sources cellulaires et notamment de cellules
d'origine humaine, ou d'autres organismes. De telles séquences homologues peuvent
être obtenues par des expériences d'hybridation. Les hybridations peuvent être
15 réalisées à partir de banque d'acides nucléiques, en utilisant comme sonde, la
séquence native ou un fragment de celle-ci, dans des conditions variables
d'hybridation (Maniatis et al.;1989).

Les séquences nucléotidiques selon l'invention peuvent être d'origine artificielle ou
20 non. Il peut s'agir de séquences génomiques, d'ADNc, d'ARN, de séquences hybrides
ou de séquences synthétiques ou semi-synthétiques. Ces séquences peuvent être
obtenues par exemple par criblage de banques d'ADN (banque d'ADNc, banque
d'ADN génomique) au moyen de sondes élaborées sur la base de séquences
présentées ci-avant. De telles banques peuvent être préparées à partir de cellules de
25 différentes origine par des techniques classiques de biologie moléculaire connues de
l'homme du métier. Les séquences nucléotidiques de l'invention peuvent également
être préparées par synthèse chimique ou encore par des méthodes mixtes incluant la
modification chimique ou enzymatique de séquences obtenues par criblage de
banques. D'une manière générale les acides nucléiques de l'invention peuvent être
30 préparés selon toute technique connue de l'homme du métier.

La présente invention a également pour objet des protéines ou des polypeptides capables d'interagir avec les présénilines. De manière préférée, les polypeptides selon l'invention son capables d'interagir avec la grande boucle de la préséniline PS-1 et/ou la grande boucle de la préséniline PS-2. Plus
5 préférentiellement, il s'agit de polypeptides capables d'interagir avec la séquence d'amino-acides 310-463 de la grande boucle de la préséniline PS-1.

Au sens de la présente invention la dénomination préséniline couvre les présénilines PS-1 et PS-2 en elles même ainsi que leurs formes homologues
10 correspondant notamment à des formes mutées de ces protéines.

La présente invention concerne également des polypeptides comprenant tout ou partie d'une séquence peptidique choisie parmi les séquences SEQ ID N°1 ou SEQ ID N°2 ou d'un dérivé de celles-ci.

15

Au sens de la présente invention, le terme séquence polypeptidique dérivée désigne toute séquence polypeptidique différant des séquences présentées en séquence SEQ ID N°1 ou SEQ ID N°2, obtenue par une ou plusieurs modifications de nature génétique et/ou chimique, et possédant la capacité d'interagir avec les
20 présénilines. Par modification de nature génétique et/ou chimique, on peut entendre toute mutation, substitution, délétion, addition et/ou modification d'un ou plusieurs résidus.

De tels dérivés peuvent être générés dans des buts différents, tels que
25 notamment celui d'augmenter leur efficacité thérapeutique ou de réduire leurs effets secondaires, ou celui de leur conférer de nouvelles propriétés pharmacocinétiques et/ou biologiques.

Le terme dérivé comprend également les séquences homologues à la
30 séquence considérée, issues d'autres sources cellulaires et notamment de cellules d'origine humaine, ou d'autres organismes, et possédant une activité de même type.

Il peut également s'agir de fragments des séquences peptidiques indiquées ci-dessus. De tels fragments peuvent être générés de différentes façons. En particulier, ils peuvent être synthétisés par voie chimique, sur la base des séquences données dans la présente demande, en utilisant les synthétiseurs peptidiques connus de l'homme du métier. Ils peuvent également être synthétisés par voie génétique, par expression dans un hôte cellulaire d'une séquence nucléotidique codant pour le peptide recherché. Dans ce cas, la séquence nucléotidique peut être préparée chimiquement en utilisant un synthétiseur d'oligonucléotides, sur la base de la séquence peptidique donnée dans la présente demande et du code génétique. La séquence nucléotidique peut également être préparée à partir des séquences données dans la présente demande, par coupures enzymatiques, ligature, clonage, etc, selon les techniques connues de l'homme du métier, ou par criblage de banques d'ADN avec des sondes élaborées à partir de ces séquences.

La présente invention a également pour objet l'utilisation des polypeptides selon l'invention, pour un ralentissement, une inhibition, une stimulation ou une modulation de l'activité des Présénilines.

En effet, il est envisageable de réguler le fonctionnement des Présénilines par le biais des protéines selon l'invention et notamment d'inhiber ou de ralentir l'activité des Présénilines mutées.

Un autre objet de la présente invention concerne un procédé de préparation des polypeptides selon l'invention selon lequel on cultive une cellule contenant une séquence nucléotidique selon l'invention, dans des conditions d'expression de ladite séquence et on récupère le polypeptide produit. Dans ce cas, la partie codant pour ledit polypeptide est généralement placée sous le contrôle de signaux permettant son expression dans un hôte cellulaire. Le choix de ces signaux (promoteurs, terminateurs, séquence leader de sécrétion, etc.) peut varier en fonction de l'hôte cellulaire utilisé. Par ailleurs, les séquences nucléotidiques de l'invention peuvent faire partie d'un vecteur qui peut être à répllication autonome ou intégratif. Plus particulièrement, des vecteurs à répllication autonome peuvent être préparés en

utilisant des séquences à réplication autonomes chez l'hôte choisi. S'agissant de vecteurs intégratifs, ceux-ci peuvent être préparés, par exemple, en utilisant des séquences homologues à certaines régions du génome de l'hôte, permettant, par recombinaison homologue, l'intégration du vecteur.

- 5 La présente invention a également pour objet des cellules hôtes transformées avec un acide nucléique comportant une séquence nucléotidique selon l'invention. Les hôtes cellulaires utilisables pour la production des peptides de l'invention par voie recombinante sont aussi bien des hôtes eucaryotes que procaryotes. Parmi les hôtes eucaryotes qui conviennent, on peut citer les cellules animales, les levures, ou
10 les champignons. En particulier, s'agissant de levures, on peut citer les levures du genre *Saccharomyces*, *Kluyveromyces*, *Pichia*, *Schwanniomyces*, ou *Hansenula*. S'agissant de cellules animales, on peut citer les cellules COS, CHO, C127, de neuroblastomes humains etc. Parmi les champignons, on peut citer plus particulièrement *Aspergillus* ssp. ou *Trichoderma* ssp. Comme hôtes procaryotes, on
15 préfère utiliser les bactéries suivantes *E.coli*, *Bacillus*, ou *Streptomyces*.

Selon un mode préféré, les cellules hôtes sont avantageusement représentées par des souches de levures recombinantes pour l'expression des acides nucléiques de l'invention ainsi que la production des protéines dérivées de ceux-ci.

- Préférentiellement, les cellules hôtes comprennent au moins une séquence ou
20 un fragment de séquence choisis parmi les séquences SEQ ID N°1 ou SEQ ID N° 2, pour la production des protéines selon l'invention.

- Une autre application des séquences d'acides nucléiques selon l'invention est la réalisation d'oligonucléotides antisens ou d'antisens génétiques utilisables comme agents pharmaceutiques. Les séquences antisens sont des oligonucléotides de petite
25 taille, complémentaires du brin codant d'un gène donné, et de ce fait capables d'hybrider spécifiquement avec l'ARNm transcrit, inhibant sa traduction en protéine. L'invention a ainsi pour objet les séquences antisens capables d'inhiber, au moins partiellement, l'expression de protéines capables d'interagir avec les présénilines. De telles séquences peuvent être constituées par tout ou partie des séquences
30 nucléotiques définies ci-avant et peuvent être obtenus par fragmentation, etc. ou par synthèse chimique.

Les séquences revendiquées peuvent être utilisées dans le cadre de thérapies géniques, pour le transfert et la production *in vivo* de séquences antisens ou l'expression de protéines ou de polypeptides capables de moduler l'activité des présénilines. A cet égard les séquences peuvent être incorporées dans des vecteurs
5 viraux ou non viraux, permettant leur administration *in vivo*. Le vecteur peut être un plasmide dont l'obtention est connue de l'homme du métier. A titre de vecteurs viraux conformes à l'invention on peut tout particulièrement citer les vecteurs de type adénovirus, rétrovirus, virus adéno-associés ou virus de l'herpès. La présente demande a également pour objet des virus recombinants défectifs comprenant une
10 séquence nucléique hétérologue codant pour un polypeptide selon l'invention.

L'invention permet également la réalisation de sondes nucléotidiques, synthétiques ou non, capables de s'hybrider avec les séquences nucléotidiques définies ci-avant ou des ARNm correspondants, utilisables dans le cadre d'une
15 thérapie génique. De telles sondes peuvent être utilisées *in vitro* comme outil de diagnostic, pour la détection des présénilines, ou encore pour la mise en évidence d'anomalies génétiques (mauvais épissage, polymorphisme, mutations ponctuelles, etc). Ces sondes peuvent également être utilisées pour la mise en évidence et
20 l'isolement de séquences d'acides nucléiques homologues codant pour des peptides tels que définis précédemment, à partir d'autres sources cellulaires et préférentiellement de cellules d'origines humaines. Les sondes de l'invention comportent généralement au moins 10 bases, et elles peuvent par exemple comporter jusqu'à l'intégralité d'une des séquences précitées ou de leur brin complémentaire. Préférentiellement, ces sondes sont, préalablement à leur utilisation, marquées. Pour
25 cela, différentes techniques connues de l'homme du métier peuvent être employées (marquage radioactif, enzymatique, etc).

Un autre objet de l'invention réside dans des anticorps ou fragment d'anticorps polyclonaux ou monoclonaux dirigés contre un polypeptide tel que défini
ci-avant. De tels anticorps peuvent être générés par des méthodes connues de l'homme du métier. En particulier ces anticorps peuvent être préparés par
30 immunisation d'un animal contre un polypeptide dont la séquence est choisie parmi les séquences SEQ ID N°1 ou SEQ ID N°2 ou tout fragment ou dérivé de celles-ci, puis prélèvement du sang et isolement des anticorps. Ces anticorps peuvent

également être générés par préparation d'hybridomes selon les techniques connues de l'homme de l'art. Les anticorps ou fragment d'anticorps selon l'invention peuvent notamment être utilisés pour inhiber et/ ou de révéler l'interaction entre les présénilines et les polypeptides tels que définis ci-avant.

- 5 Un autre objet de la présente invention concerne un procédé d'identification de composés capables de moduler ou d'inhiber totalement ou partiellement l'interaction entre les présénilines et les polypeptides selon l'invention.

La mise en évidence et/ou l'isolement de modulateurs ou de ligands des polypeptides selon l'invention, est réalisé selon les étapes suivantes :

- 10 - on met en contact une molécule ou un mélange contenant différentes molécules, éventuellement non-identifiées, avec une cellule recombinée telle que décrite ci-dessus exprimant un polypeptide de l'invention dans des conditions permettant l'interaction entre ledit polypeptide et ladite molécule dans le cas où celle-ci posséderait une affinité pour ledit polypeptide, et,
- 15 - on détecte et/ou isole les molécules liées au dit polypeptide de l'invention.

Dans un mode particulier, ce procédé de l'invention est adapté à la mise en évidence et/ou l'isolement d'agonistes et d'antagonistes de l'interaction entre les présénilines et les polypeptides de l'invention.

- 20 Un autre objet de l'invention concerne l'utilisation d'un ligand ou d'un modulateur identifié et/ou obtenu selon le procédé décrit ci-avant comme médicament. De tels ligands ou modulateurs peuvent en effet permettre de traiter certaines affections neurologiques.

L'invention a encore pour objet toute composition pharmaceutique comprenant comme principe actif au moins un polypeptide tel que défini ci-avant.

Elle a aussi pour objet toute composition pharmaceutique comprenant comme principe actif au moins un anticorps et/ou un fragment d'anticorps tel que
5 défini ci-avant, et/ou un oligonucléotide antisens, ainsi que toute composition pharmaceutique comprenant comme principe actif au moins une séquence nucléotidique telle que définie ci-avant.

Par ailleurs, elle a aussi pour objet les compositions pharmaceutiques dans lesquelles les peptides, anticorps et séquences nucléotidiques définis ci-avant ou des
10 fragments de ceux-ci sont associés entre-eux ou avec d'autres principes actifs.

Elle a également pour objet les compositions dans lesquelles les séquences nucléotidiques selon l'invention sont incorporées dans un vecteur recombinant viral ou non viral.

Les compositions pharmaceutiques selon l'invention peuvent être utilisées
15 pour moduler l'activité des présénilines. Plus préférentiellement, ces compositions sont destinées à ralentir ou inhiber au moins partiellement l'activité des présénilines et notamment les formes mutées. Il s'agit plus préférentiellement de compositions pharmaceutiques destinées au traitement de maladies neurodégénératives comme par exemple la maladie d'Alzheimer et la trisomie 21.

20

Liste des figures

Figure 1. Schéma de la préséniline I et de la partie utilisée pour le test en double
25 hybride.

Figure 2. Isolation de souches yBM-C et yBM-D par test d'auxotrophie des levures portant différents plasmides contenant notamment le clone C (correspondant à la SEQ ID N°1) et le clone D (correspondant à la SEQ ID N°2)

30

Figure 3. Expression des ARNm correspondants aux peptides C et D dans les tissus humains

Figure 4. Schéma du vecteur d'expression mammifère utilisé.

Figure 5. Expression des protéines de fusion C et D en cellules mammifères.

Figure 6 (a, b, c). Cartes des principaux vecteurs

5

MATERIELS ET METHODES

1) Les souches de levures utilisées sont les suivantes:

10 *S.cerevisiae* YCM 79 (*MATa, ade2, trp1, leu2, lys2, ura3, his3, gal4, gal80, met16 :: GAL1/10-URA3 lacZ, HIS3 :: GAL7-phleo^R*).

S.cerevisiae PCY 2 (*MATa, Dgal4, Dgal80, ade2, trp1, leu2, lys2, his3 URA3 :: GAL1/10- lacZ*).

S.cerevisiae HF7 C (*MATa, ura3, his3, ade2, trp1, leu2, lys2, gal4, gal80, can^R, URA3 :: GAL4 binding site- CYC1-lacZ, LYS2 :: GAL1- HIS3*).

15 *S.cerevisiae* L 40 (*MATa, his3, ade2, trp1, leu2, lys2, URA3 :: LexA-lacZ, LYS2 :: LexA-HIS3*).

2) Les souches de bactéries utilisées sont les suivantes :

E.coli TG1 (*F'[traD 36, lacI^q, lacZ D M15, proA⁺ B⁺], supE, thi-1, D (lac-pro AB), D (hsdM-mcrB)5 (r_K-m_K- McrB-)*).

20 *E.coli* BL21 (DE3) (*F-, ompT, hsdSB (r_B⁻ m_B⁻), gal, dcm, (DE3)*).

3) Les milieux de culture mis en oeuvre pour les diverses souches sont les suivants :

milieu YPD complet pour les levures :

-Bacto-yeast extract (10 g/l) (Difco)

25 -Bacto-peptone (20 g/l) (Difco)

-Glucose (20 g/l) (Merk)

Ce milieu peut être solidifié par addition d'agar (20 g/l) et rendu stérile à 110 °C pendant 20 min dans un autoclave.

milieu YNB minimum pour les levures :

5 -Bacto-yeast nitrogen base without amino acids (6,7 g/l) (Difco)

-Glucose (20g/l) (Merk)

Ce milieu peut être solidifié par addition d'agar (20g/l) et stérilisée par autoclavage. Les acides aminés ou les bases azotées nécessaires à la croissance des levures auxotrophes sont ajoutés ensuite à 50mg/l. On ajoute aussi de l'Ampicilline à
10 100µg/ml pour éviter les contaminations bactériennes.

milieu LBcomplet pour les bactéries :

-Bacto-yeast extract (5g/l) (Difco)

-Bacto-tryptone (10 g/l) (Difco)

-NaCl (5 g/l) (Difco)

15 Ce milieu peut être solidifié par addition d'agar (12 g/l) et stérilisé par autoclave. L'ampicilline ou la kanamycine sont ajoutées à 100 µg/ml pour sélectionner les bactéries ayant été transformées par les plasmides portant les marqueurs de résistance correspondants.

4) Les plasmides utilisés sont les suivants :

20 Le vecteur pGBT9 (fig. 6 (a)), est un plasmide navette de 5,5 kb pouvant être propagé et sélectionné dans les bactéries *E.coli* aussi bien que dans les levures. Il possède une origine de répllication ColE1 et un gène de résistance à l'ampicilline (propagation et sélection dans *E.coli.*) ainsi qu'une origine de répllication 2mm et un

gène TRP1 (sélection chez les levures *trp1*- déficientes pour la synthèse du tryptophane). Ce plasmide est utilisé pour exprimer chez la levure des protéines de fusion entre le domaine de liaison à l'ADN (DLA) de GAL4 et des peptides d'intérêt. Il possède un site multiple de clonage en aval de la région DLA de GAL4, le tout
5 étant situé entre le promoteur et le terminateur du gène ADH1 codant pour l'alcool deshydrogénase 1. Le DLA de GAL4 contient une séquence signal de localisation nucléaire.

Les séquences correspondant à différentes parties de PS-I ont été insérées aux sites de clonage EcoRI/Sall.

- 10 Le vecteur pGAD10 (fig. 6 (a)), est un plasmide navette de 6,6 kb. Il possède une origine de répllication ColE1 et un gène de résistance à l'ampicilline qui permettent sa propagation et sa sélection dans *E.coli*.. Il contient aussi une origine de répllication de levure (2mm) et un gène de sélection LEU2 pour isoler les levures transformées. Il possède un site multiple de clonage en aval de la séquence codant pour le domaine
15 d'activation (DA) de GAL4 entre le promoteur et le terminateur de ADH1. Le DA de GAL4 contient une séquence signal de localisation nucléaire.

Ce plasmide a été utilisé pour réaliser une banque d'ADNc de cerveau humain commerciale (Clontech). Les séquences d'ADNc ont été insérées au site de clonage EcoRI.

- 20 Le vecteur pLex 9 (fig. 6 (b)), est un plasmide navette de 5 kb. Comme pGBT9, il possède l'origine ColE1, le gène Amp^R, ainsi que l'origine de répllication 2mm et le gène TRP1. Cependant, la séquence codant pour le domaine de liaison à l'ADN est celle du répresseur bactérien Lex A. Le promoteur et le terminateur servant à l'expression de la protéine de fusion sont également ceux du gène ADH1. Le DLA de
25 Lex A contient la séquence signal de localisation nucléaire de SV40.

Les séquences correspondantes à la grande boucle de PS-I ont été insérées aux sites de clonage EcoRI/Sall.

Le vecteur pGEX-4T1 (fig. 6 (b)), est un plasmide bactérien de 4,9 kb. Il possède une origine de répllication pBR322, un gène de résistance à l'ampicilline, un gène lacIq pour réprimer le système *beta*-galactosidase. Ce plasmide est utilisé pour exprimer chez la bactérie des protéines de fusion entre la protéine glutathion S-transferase (GST) et des peptides d'intérêt. Il possède un site multiple de clonage en aval du gène de la GST pour insérer les séquences codant pour les peptides. L'ensemble est contrôlé par le promoteur tac à haut niveau d'expression inductible par l'IPTG. Après production de la protéine, le peptide d'intérêt peut être séparé de la GST par clivage à la thrombine, car la séquence reconnue et clivée par cette enzyme a été introduite entre ces deux régions.

Le vecteur pET-29a (fig. 6 (c)), est un plasmide bactérien de 5,4 kb. Il possède une origine de répllication pBR322, une origine de répllication du phage f1, un gène de résistance à la kanamycine, un gène lacIq. Ce plasmide est utilisé pour exprimer chez la bactérie des peptides d'intérêt fusionnés à une séquence S-Tag (partie catalytique de la protéine S ribonucléase) et une séquence His-Tag (peptide de 10 histidines) qui facilitent la détection et la purification de la protéine de fusion. Le promoteur de transcription T7 permet une forte expression après induction à l'IPTG. Une séquence reconnue et clivée par la thrombine a été introduite entre les séquences S-Tag et His-Tag.

5) Les oligonucléotides de synthèse employés :

A partir de la séquence de PS-I (Sherrington et al., 1995), différents couples d'oligonucléotides on été définis afin d'obtenir les fragments d'ADN correspondant aux régions hydrophiles de PS-I. Ces oligonucléotides permettent d'introduire un site EcoRI en 5' et un site SalI en 3' de la séquence après amplification par PCR.

oligoA CAA GAA TTC TGT CCG AAA GGT CCA CTT (SEQ ID N° 3)

oligoB CAA GTC GAC TCA GGT TGT GTT CCA GTC TCC (SEQ ID N°

4)

Les oligonucléotides A et B ont servi à obtenir les fragments BL6 et BM issus de la grande boucle.

oligoC CAA GAA TTC TCA ACA ATG GTG TGG TTG (SEQ ID N° 5)

oligoD CAA GTC GAC TCA TTC CTC TGG GTC TTC ACC (SEQ ID N° 6)

Les oligonucléotides C et D ont servi à obtenir le fragment BR issu d'une région plus courte de la grande boucle.

oligoE CAA GAA TTC ACA TTA TTA CTC CTT GCC (SEQ ID N° 7)

oligoF CAA GTC GAC TCA GAT ATA AAA TTG ATG CAA (SEQ ID N° 8)

10 Les oligonucléotides E et F ont servi à obtenir le fragment Ct, issu de la partie C-terminale de PSI.

oligoG CAA GAA TTC ATG ACA GAG TTA CCT GCA (SEQ ID N° 9)

oligoH AA GTC GAC TCA ATG CTT GGC GCC ATA TTT (SEQ ID N° 10)

15 Les oligonucléotides G et H ont servi à obtenir le fragment Nt, issu de la partie N-terminale de PSI.

Les oligonucléotides I, J et K ont été utilisés pour amplifier et séquencer les inserts des plasmides pGBT9 et pGAD10. Ils sont, respectivement, complémentaires des séquences du DLA de GAL4, du DA de GAL4 et du terminateur ADH1.

oligoI GCG ACA TCA TCA TCG GAA GAG AG (SEQ ID N° 11)

20 oligoJ CTA TTC GAT GAT GAA GAT ACC CC (SEQ ID N° 12)

oligoK GGC GAA GAA GTC CAA AGC (SEQ ID N° 13)

6) Les méthodes de biologie moléculaire.

Préparation de l'ADN plasmidique.

Les préparations en petite quantité et en grande quantité d'ADN plasmidique ont été effectuées selon les protocoles décrits dans Maniatis *et al.* (1989).

Amplification en chaîne par la polymérase Taq thermostable (ou PCR).

- 5 La PCR permet d'amplifier une séquence d'ADN entre deux régions connues où peuvent se fixer des oligonucléotides (de 20 pb environ) qui serviront d'amorces à une polymérase thermostable. Les réactions sont effectuées dans un volume final de 100 µl en présence de la matrice d'ADN (50 ng), de dNTP (0,2 mM), de tampon PCR (TrisHCl pH 8,5, 10 mM, MgCl₂ 1 mM, KCl 5 mM, gélatine 0,01%), des deux
10 oligonucléotides (500 ng chacun) et de 2,5 UI d'Ampli Taq DNA polymérase. Le mélange est recouvert de 2 gouttes d'huile de paraffine pour limiter l'évaporation de l'échantillon.

- Le programme utilisé comporte 32 cycles : 1 cycle de dénaturation de la matrice (5 min à 95 °C), 30 cycles (1 min de dénaturation à 95 °C, 1min d'hybridation des
15 oligonucléotides sur la matrice d'ADN à 50 °C, 1 min d'élongation par la Taq polymérase à 72 °C) enfin 1 cycle d'élongation finale (5 min à 72 °C).

La PCR peut être effectuée directement sur des cultures de bactéries ou de levures en phase exponentielle de croissance.

Séquençage par fluorescence.

- 20 La technique de séquençage utilisée est dérivée de la méthode de Sanger *et al.* (1977) et adapté pour le séquençage par fluorescence développé par Applied Biosystems. Le protocole utilisé est celui décrit par les concepteurs du système (Perkin Elmer, 1995).

Insertion d'un fragment d'ADN dans un plasmide par ligation.

- L'insert provenant d'un ADN plasmidique ou d'une réaction de PCR et le plasmide
25 vecteur sont digérés 1 heure par des enzymes de restriction à 1 U/µg d'ADN dans

leur tampon respectif (Biolabs). Les enzymes sont choisies en fonction des sites de clonage pour introduire l'insert en phase avec les autres domaines du vecteur.

Les fragments issus de la digestion sont séparés selon leur taille par électrophorèse sur un gel " préparatif " d'agarose 0,8% fait dans du TBE (Tris 90
5 mM, borate 90 mM, EDTA 2mM, pH 8) avec 0,5 µg/ml de BET. Du bleu de dépôt (1/10 de volume) est ajouté aux échantillons avant leur chargement sur le gel à côté d'un marqueur de poids moléculaire (1 Kb ladder, Gibco BRL). La migration s'effectue à 90 volts/cm constants dans du TBE.

L'ADN est révélé sous U.V. par la fluorescence du BET qui s'est intercalé
10 entre les bases. Les morceaux de gel contenant les fragments d'ADN de tailles voulues (vecteur et insert) sont découpés au scalpel, introduits dans un boudin à dialyse avec du TBE et soumis à une électrophorèse pendant 15 min à 100 volts. L'ADN extrait du gel est ensuite récupéré, traité par un volume de phénol/chloroforme/isoamylate (25/24/1), précipité par 3 volumes d'éthanol absolu
15 en présence de NaCl 0,25 M, lavé une fois à l'éthanol 70%, séché et enfin repris dans 20µl d'H₂O.

Après avoir estimé sur gel les quantités respectives de vecteur et d'insert, ces derniers sont mélangés dans un rapport de 1/1 en présence de 40 U.I. de T4 DNA ligase, du tampon de ligation 1X final (TrisHCl 50 mM, pH 7,5, MgCl₂ 10 mM,
20 DTT 10 mM, ATP 1 mM) dans un volume total de 20 µl. La ligation s'effectue à 20 °C pendant au moins une heure.

Transformation des bactéries par un plasmide.

1) La méthode dite au " TSB " (d'après Chung et Miller, N.A.R. vol 16, 1988) consiste à fragiliser la paroi des bactéries par un traitement au DMSO pour permettre
25 l'entrée de l'ADN dans les cellules. Son efficacité est de 10⁶ transformants/µg d'ADN. Les bactéries sont cultivées dans un milieu LB liquide pendant environ 3 heures sous agitation (jusqu'à A₆₀₀=0,6). La culture est centrifugée 10 min à 3000

rpm, le culot est repris dans 1/10 du volume initial par du TSB (milieu LB, PEG₄₀₀₀ 10%, DMSO 5%, MgCl₂ 10 mM, MgSO₄ 10 mM), et incubé 15 min à 4°C. Environ 50 ng d'ADN (10 µl pour le produit de ligation) sont ajoutés à 100µl de bactéries compétentes, puis le mélange est incubé 30 min à 4 °C. Après addition de TSBglu
5 (TSB, glucose 20 mM) et incubation à 37 °C pendant 45 min, les bactéries sont étalées sur milieux LB avec l'antibiotique correspondant au gène de résistance porté par le plasmide.

2) L'électroporation est une technique 1000 fois plus efficace que la méthode au TSB, elle est utilisée lorsqu'on dispose de très faibles quantités d'ADN (comme après
10 l'extraction des plasmides contenus dans les levures). Elle consiste à fragiliser la paroi bactérienne par un choc électrique pour introduire l'ADN dans les cellules.

Les bactéries sont cultivées dans un milieu LB liquide sous agitation jusqu'à $A_{600}=0,6$. La culture, incubée préalablement 15 min à 4 °C, est centrifugée 10 min à 3000 rpm et le culot est lavé successivement par 1 volume d'H₂O glacée, 1/2 volume
15 d'H₂O glacée, et 1/5 volume de glycérol 10% glacé. Entre chaque lavage, les bactéries sont centrifugées 10 min à 3000 rpm pour enfin être reprises dans 1/1000 volume de glycérol 10% glacé. 40 µl de cette suspension de bactéries sont mélangés avec 1µl d'ADN dans une cuve à électroporation. Les cuves sont placées dans le Gene Pulser BioRad pour que les bactéries subissent un choc électrique (25 mF,
20 2,5kV, 200 W) . Après incubation à 37 °C pendant 45 min, les bactéries sont étalées sur milieu LB contenant l'antibiotique servant à la sélection.

Hybridation des ARNm après transfert sur membrane (ou Northern Blot).

Les membranes utilisées sont des membranes de nitrocellulose commerciales Clontech où les ARNm issus de différents tissus humain ont déjà été transférés et
25 fixés.

La sonde d'ADN radioactive est préparée suivant le protocole « Rediprime DNA Labelling System » Amersham. Ce protocole permet d'obtenir un taux

d'incorporation du [^{32}P] dCTP (50 mCi) supérieur à 55 % dans le fragment d'ADN après une incubation de 10 min à 37 °C et de détecter une bande d'ARN équivalent à 0,5 pg après une hybridation de 2 heures et une nuit de mise en autoradiographie.

Les membranes sont préhybridées avec 8 ml de solution d'hybridation (NaCl 1 M, NaH₂PO₄ 10 mM pH 7,4 ; Ficoll 10 mg/ml, polyvinylpyrrolidone 10 mg/ml, BSA 10 mg/ml, ADN de sperme de saumon 100 mg/ml, SDS 2%) environ 2 heures à 42 °C sous agitation, puis elles sont hybridées avec la sonde à $2 \cdot 10^6$ cpm/ml mélangée avec 8 ml de solution d'hybridation pendant 24 heures à 42 °C sous agitation, elles sont ensuite rincées plusieurs fois avec la solution 1 de lavage (NaCl 0,3 M, sodium citrate 30 mM pH 7, SDS 0,05%), lavées pendant 40 min dans cette même solution, incubées 40 min à 50 °C dans la solution 2 de lavage (NaCl 15 mM, sodium citrate 1,5 mM pH 7, SDS 0,1%) et enfin elles sont soumises à une autoradiographie.

7) Les méthodes pour le système double hybride.

Principe du système.

Le système double hybride, développé par Song et Fields depuis 1989, est une technique de clonage par interaction *in vivo*, chez la levure *Saccharomyces cerevisiae*. Il repose sur le fait que certains transactivateurs eucaryotes ne nécessitent que deux domaines pour fonctionner: un domaine de liaison à l'ADN (DLA) qui se fixe sur une séquence spécifique et un domaine d'activation (DA) qui recrute la machinerie de transcription (ARN polII). Lorsque les deux domaines sont séparés, le DA ne peut pas se positionner correctement et la transcription n'est pas induite.

Dans ce système, le DLA est fusionné à une protéine X et le DA à une protéine Y. Si les protéines X et Y interagissent, un complexe transactivateur fonctionnel se forme et induit l'expression d'un gène rapporteur placé en aval de la séquence reconnue par le DLA. L'expression du gène permet de révéler indirectement l'interaction entre les deux protéines.

Transformation de la levure par un ou deux plasmides différents.

Les levures sont rendues compétentes par un traitement au LiAc/PEG d'après la méthode décrite par Gietz (Gietz *et al.*, 1995).

Les levures sont cultivées sur milieu solide YPD à 28 °C pendant une nuit. Elles sont
5 ensuite resuspendues dans 1 ml de solution I stérile (LiAc 0,1 M, TrisHCl 10 mM, EDTA 1mM pH7,5), centrifugées 1 min à 3000 rpm et reprises à nouveau dans 1 ml de solution I. 50 µl de la suspension sont alors mélangés à 5 µg d'ADN de sperme de saumon (ADN entraîneur), 1 à 5 µg d'ADN plasmidique et 300 µl de solution II stérile (LiAc 0,1 M, TrisHCl 10 mM, EDTA 1 mM pH 7,5, PEG₄₀₀₀ 50%). Le
10 mélange est incubé 30 min à 28 °C puis subit un choc thermique pendant 20 min à 42 °C. Après centrifugation 1 min à 3000 rpm, les cellules sont lavées une fois à l'eau stérile et reprises dans 100 µl d'eau stérile. Puis, les levures sont étalées sur milieu gélosé YNB additionné des acides aminés et bases azotées nécessaires à leur
15 croissance. Pour permettre la sélection des souches de levure transformées, les acides aminés et les bases azotées correspondant aux marqueurs de sélection portés par les plasmides ne sont pas ajoutés. Les boîtes sont mises à l'étuve à 28 °C pendant 3 jours.

Transformation de la levure par une banque d'ADNc.

La banque d'ADNc utilisée est une banque commerciale (Clontech) constituée à
20 partir d'ARNm d'un cerveau humain normal (homme de 37 ans décédé à la suite d'un traumatisme cérébral). Les ADNc ont été clonés dans le plasmide pGAD10 en aval de la région du DA de GAL4 au site EcoRI. La moyenne de la taille des inserts est de 1,4 kb avec un écart type de 1 kb. La banque a été amplifiée une fois dans *E.coli* DH5a pour obtenir $1,2 \times 10^6$ clones indépendants dont 80% contiennent un
25 insert.

Pour préserver une bonne représentativité de la banque, il est nécessaire d'avoir un nombre de transformants deux à dix fois plus grand que celui des clones indépendants

obtenus lors de la construction de la banque. Nous utilisons donc un protocole qui permet d'obtenir une bonne efficacité de transformation (environ 10^7 transformants/100 μ g d'ADN).

Les levures YCM79 préalablement transformées par le plasmide pGBT9 contenant
5 une séquence correspondant à une région de PS-I sont cultivées une nuit dans 50 ml de milieu YNB +His +Lys +Ade +Met +Leu +Ura à 28 °C sous agitation. La densité optique de cette préculture est mesurée pour évaluer le nombre de cellules par ml et permettre le calcul du volume à inoculer dans 250 ml de milieu YPD afin d'obtenir une culture de 10^7 cellules/ml (environ $A_{600}=0,8$) après une nuit à 28 °C sous
10 agitation.

Les cellules sont ensuite récoltées par centrifugation à 3000 rpm pendant 10 min, lavées deux fois avec 100 ml d'H₂O stérile, reprises dans 10 ml de solution I de transformation. La suspension est centrifugée 10min à 3000rpm et les cellules resuspendues dans 2,5 ml de cette même solution. Une partie de la suspension est
15 prélevée (200 μ l): la moitié servira de témoin négatif non transformé, l'autre moitié sera transformée par un plasmide contrôle dont l'efficacité de transformation est connue. La suspension restante est mélangée à 100 μ l de banque plasmidique d'ADNc à 1mg/ml et 20 ml de solution II. Le mélange est incubé 1 heure à 28 °C sous agitation puis 20 min à 42 °C avant d'être centrifugé 15 min à 3000 rpm. Le
20 culot cellulaire est lavé une fois à l'eau et deux fois au PBS afin de bien éliminer le PEG qui a un effet toxique sur les levures, puis est repris dans 2,5 ml de PBS stérile (Na₂HPO₄ 50 mM, KCl 10 mM, MgSO₄ 1 mM, pH 7).

250 ml de milieu YNB +His +Lys +Ade +Met +Ura sontensemencés par cette suspension et incubés 4 heures à 28 °C sous agitation pour permettre l'expression des
25 protéines hybrides, leur interaction éventuelle et l'expression du gène rapporteur *URA3*. Cependant, cette étape a pour conséquence de multiplier le nombre de clones positifs. Pour évaluer le taux d'amplification du nombre de colonies, 1 ml de culture est prélevé avant et après cette incubation et différentes dilutions sont étalées sur

milieu gélosé YNB +His +Lys +Ade +Met +Ura. Après croissance, le nombre de clones Leu+ Trp+ est déterminé pour évaluer la fréquence de transformation avant et après la période d'incubation et le taux d'amplification en est déduit.

Les cellules sont alors centrifugées 10 min à 3000 rpm, lavées deux fois à l'eau stérile, reprises dans 10 ml d'YNB et étalées sur 10 boîtes de 435 cm² contenant chacune 200 ml de milieu sélectif YNB +His +Lys +Ade +Met. Les boîtes sont mises à l'étuve à 28 °C pendant 3 jours.

Extraction de l'ADN plasmidique des levures.

Cette technique extrait aussi bien les plasmides que l'ADN génomique contenu dans les levures. Pour isoler les plasmides, il suffit de transformer des bactéries où seul l'ADN plasmidique peut se répliquer et d'analyser chaque clone après minipréparation. Les levures contenant les plasmides sont cultivées sur milieu solide YNB sélectif à 28 °C pendant une nuit. Elles sont ensuite resuspendues dans 200 ml de solution de cassage (Triton X100 2%, SDS 1%, NaCl 100 mM, Tris 10 mM pH 8, EDTA 1 mM) en présence de billes de verre (450 µm de diamètre) et de 200 ml de phénol/chloroforme. Le mélange est agité 15 min sur un vortex puis centrifugé 2 min à 14000 rpm. La phase aqueuse est dialysée 1 heure sur membrane. Cette solution peut servir à la transformation des bactéries par électroporation.

Révélation de l'activité *beta*-galactosidase.

Ce test permet de vérifier l'expression du gène rapporteur *LacZ* codant pour la β -galactosidase dans les levures transformées.

Une feuille de nitrocellulose est déposée sur les clones de levures individualisés. La réplique est plongée 3 fois pendant 30 secondes dans de l'azote liquide pour casser les levures et permettre l'accès du substrat à la *beta*-galactosidase. La feuille est ensuite posée, colonies vers le haut, sur un papier Whatman imbibé d'une solution de PBS/Xgal préparée extemporanément (100 µl d'une solution de substrat Xgal à 40

mg/ml en diméthylformamide dans 5 ml de PBS) puis placée à 37 °C, température optimale pour l'activité enzymatique testée.

Le temps d'apparition de la couleur bleue qui révèle l'activité β -gal est très variable, de quelques minutes à plusieurs heures. Le test s'effectue en présence d'un témoin
5 positif d'interaction qui vire rapidement au bleu et d'un témoin négatif qui reste blanc.

Test en goutte.

Ce test permet d'analyser le phénotype d'une souche de levure. Une goutte (environ 5 ml) d'une suspension de levures légèrement trouble est déposée sur différents milieux
10 gélosés sélectifs et non sélectifs, et la croissance des levures est observée après 2 jours d'incubation à 28 °C.

Ségrégation des plasmides.

Il s'agit de tester par une ségrégation plasmidique si le phénotype des levures transformées est bien lié à la présence des plasmides.

15 Les levures sont cultivées une nuit sur milieu complet YPD à 28 °C. Ce milieu non sélectif permet la croissance des levures conservant ou ne conservant pas les plasmides après la division cellulaire. Différentes dilutions de cette culture sont étalées pour obtenir entre 50 et 100 clones par boîte de milieu YPD après une nuit
20 d'incubation à 28 °C. Le phénotype des clones est alors testé par répliques sur velours sur différents milieux sélectifs et la fréquence des cellules ne contenant pas de plasmide peut être estimée après 2 jours d'incubation à 28 °C.

8) Les méthodes pour le test d'interaction *in vitro* .

Principe du test.

Il s'agit de vérifier l'interaction entre deux protéines en utilisant une technique
25 biochimique de chromatographie d'affinité. Un extrait cellulaire contenant un des

partenaires supposés de la protéine est ajouté à un support où la protéine appât a été préalablement fixée. Si l'interaction existe entre la protéine et son éventuel partenaire, celui-ci aussi, sera retenu sur le support.

Expression des protéines de fusion chez *E.coli* .

- 5 Les protéines de fusion sont produites dans *E.coli* BL21(DE3) par l'intermédiaire des vecteurs d'expression pET et pGEX inducible par l'IPTG (Studier, 1991).

Les bactéries BL21(DE3), préalablement transformées par les plasmides recombinants, sont cultivées une nuit sous agitation dans du milieu LB avec l'antibiotique servant à la sélection (ampicilline pour pGEX ou kanamycine pour pET
10 à 100µg/ml). 5 ml de cette préculture sont inoculés dans 45 ml de milieu LB+antibiotique afin d'obtenir une culture à $A_{600}=0,6$ après 1 heure d'agitation à 37 °C. La production des protéines est alors induite par addition de 0,1 mM d'IPTG et les bactéries sont remises à 37 °C sous agitation entre 1 heure et 4 heures. Les
cellules exprimant les protéines sont récoltées par centrifugation à 6000 rpm pendant
15 5 min. Les culots peuvent être gardés à -20 °C.

Extraction et analyse des protéines produites chez la bactérie.

Les culots bactériens contenant les protéines de fusion sont resuspendus dans 5 ml de TE (TrisHCl 50 mM pH 8, EDTA 2 mM). Les cellules sont lysées par du lysozyme 0,1 mg/ml final et 1% Triton X-100. Elles sont incubées 15 min à 30 °C avant d'être
20 placées dans la glace pour être soumises aux ultra-sons jusqu'à ce que la suspension bactérienne devienne fluide. Le lysat bactérien est centrifugé 15 min à 10000 rpm à 4 °C. Le surnageant contenant la protéine de fusion solubilisée peut être conservé à -20 °C.

A cette étape, les différentes fractions obtenues sont analysées par électrophorèse sur
25 gel dénaturant (SDS PAGE). 60 µl de la suspension bactérienne avant la centrifugation (fraction totale), 60 µl de surnageant (fraction soluble) ainsi que 60 µl

du culot resuspendu dans 5 ml de TE (fraction insoluble) sont prélevés. Les protéines contenues dans les différentes fractions sont dénaturées 10 min à 95 °C en présence de 1/4 de volume de bleu de dépôt 4xL (TrisHCl 0,2 M pH 6,8, SDS 5%, glycérol 5%, 2-mercaptoethanol 0,5%, bleu de bromophénol) et chargées sur un gel de SDS polyacrylamide 12% où elles sont séparées en fonction de leur masse moléculaire par electrophorèse. Après migration à 125 volts/cm, les protéines sont détectées par coloration du gel au bleu de Coomassie.

Solubilisation des protéines.

Lors de l'expression d'un gène exogène dans E.coli peuvent apparaître des agrégats de la protéine exogène formant des corps d'inclusion. La protéine recombinante est alors facile à purifier par centrifugation mais elle est souvent difficile à solubiliser.

Les protéines se trouvant dans les fractions insolubles sont dénaturées par de l'urée 6 M et du DTT 10 mM pendant 10 min à 37 °C puis renaturées par dialyse dans trois tampon successifs pendant au moins 6 heures chacun (TrisHCl 5 mM pH 8 / urée 2 M; TrisHCl 5 mM, pH 8 / urée 1 M; TrisHCl 5 mM, pH 8). Après 15 min de centrifugation à 10000 rpm, environ 5 à 10% de la protéine se trouve dans le surnageant qui peut être conservé à -20 °C.

Chromatographie d'affinité.

Pour que la rétention des protéines s'opère dans de bonnes conditions, la concentration protéique des échantillons doit être supérieure à 40 mg/ml. Cette concentration peut être estimée par une méthode colorimétrique (réactif de Bradford, par exemple).

200 µl de fraction soluble contenant la GST non fusionnée ou la GST-BM (boucle mutée de la préséniline 1) sont incubés 2 min à 20 °C en présence de 300 µl de « résine glutathion ». La fraction non retenue par les billes est prélevée. La résine est lavée 4 fois par du tampon de lavage (TrisHCl pH 8, 20 mM, NaCl 150 mM, Triton

X100 0,1%) afin d'éliminer les protéines fixées par adsorption non spécifique. 150 µl de résine ayant liée la GST ou la GST-BM sont incubés 1 heure à 4 °C sous agitation avec 600 µl de fraction soluble contenant STag ou STag-Pep126 . Les billes d'agarose sont ensuite centrifugées 1 min à 3000 rpm et lavées 4 fois par du tampon
5 de lavage après prélèvement de la fraction non retenue.

Un aliquote de chaque fraction (soluble, non retenue et retenue par les billes) est déposée sur gel SDS PAGE après dénaturation à 95 °C en présence de bleu de dépôt et séparées par électrophorèse. Les protéines sont, ensuite, transférées sur membrane de nitrocellulose puis la membrane est colorée au rouge Ponceau (20 µg/ml) pour
10 vérifier que le transfert des protéines a bien été réalisé.

Pour pouvoir détecter spécifiquement les protéines fusionnées avec le S-Tag, la membrane de nitrocellulose est saturée avec 1% de gélatine dans du TBST (TrisHCl 10 mM pH 8, NaCl 150 mM, 0,1% Tween20) pendant 15 min et incubée 15 min avec la protéine S liée à la phosphatase alcaline. Après 4 lavages avec le TBST, l'activité
15 de la phosphatase alcaline est révélée par addition d'un tampon PA 1X contenant du NBT et du BCIP selon la méthode Novagen.

9) Les méthodes d'expression des protéines recombinantes en cellules mammifères.

20 Construction d'un vecteur d'expression en cellules mammifères

Pour l'expression en cellules mammifères, les séquences ID N° 1 et ID N° 2 ont été sousclonées par des méthodes classiques dans un vecteur sous le contrôle d'un promoteur viral fort (CMV).

Transfection de cellules

La méthode établie pour les cellules COS1 ou les cellules CHO, consiste à utiliser un lipofectant dans un ratio de 1 pour 8 (poids/poids) par-rapport à l'ADN et un peptide synthétique (H1) au même ratio afin d'optimiser la compaction de l'ADN et l'efficacité de transfection. Cette méthode repose notamment sur la neutralisation des charges des phosphates de l'ADN par les charges positives du lipofectant.

Les cellules COS1 sont cultivées en incubateur à 37 °C, 95% d'humidité et 5 % de CO₂ dans le milieu DMEM (Dulbecco's Modified Eagle's Medium) contenant 4.5 g/l glucose (Gibco-BRL) supplémenté avec 3 % de L-Glutamine, 1% de pénicilline-streptomycine et 10 % de Sérum de Veau Foetal.

La veille de la transfection, les cellules sontensemencées à une densité de $2.5 \cdot 10^6$ cellules par boîte de 100 mm. Le jour de la transfection les cellules sont rincées 2 fois par du PBS (Phosphate Buffer Saline) et une fois en OptiMEM (composition brevetée; Gibco-BRL) pour une habitude d'au moins 15 minutes en incubateur.

Par équivalent-boîte de 100 mm, 8 µg d'ADN plasmidique au total sont ajoutés à 300 µl d'OptiMEM et 64 µg de peptide synthétique (H1). Après avoir vortexé vigoureusement pendant 10 secondes, la lipofectamine (32 µl, soit 64 µg) diluée dans 300 µl d'OptiMEM est ajoutée au mélange précédent, après 5 minutes d'attente. L'ensemble est une nouvelle fois vortexé vigoureusement puis laissé 30 minutes reposer. Cinq millilitres d'OptiMEM sont ajoutés par tube et le mélange vortexé est placé sur les cellules (dont on a au préalable aspiré le milieu). Les cellules sont alors placées en incubateur pendant 4 heures, au terme desquelles le mélange est remplacé par du milieu complet.

Lyse des cellules et dosage des protéines

Les cellules sont le plus souvent lysées 48 heures après la transfection (au maximum habituel d'expression). Le tampon de lyse contient 10 mM de Tris pH 7.5, 1 mM d'EDTA, 1 % de Triton X100, 1 % de NP40 et un cocktail d'inhibiteurs de protéases (Complete©, Boehringer-Mannheim). Pour chaque plaque, après un rinçage au PBS,

800 µl du tampon froid sont ajoutés. Les lysats subissent alors une sonication suivie d'une agitation au barreau magnétique à 4 °C pendant une nuit. Une centrifugation de 30 minutes à 15000 x g sépare le culot du surnageant.

Les protéines solubles sont alors dosées selon le kit BCA (Pierce) afin de pouvoir
5 normaliser les expériences suivantes.

Immunotransfert

Les échantillons (lysats de cellules) sont dénaturés dans un volume égal de tampon de dépôt (125 mM Tris pH 6.8, 4% m/v SDS, 20 % glycérol, 0.02 % Bromophénol Blue, 50 mM Dithiothreitol) à 95 °C pendant 5 minutes. Pour l'analyse de
10 l'expression des présénilines, les échantillons sont dénaturés en présence de 8 M urée et à 37 °C afin d'éviter l'agrégation propre aux présénilines à 95 °C.

Les échantillons sont déposés sur des gels Tris-Glycine (Novex), avec un pourcentage d'acrylamide différent selon le poids moléculaire à discriminer. Un marqueur de poids moléculaire est également déposé (Broad Range, BioRad). La migration a lieu
15 pendant environ 2 heures à 100 volts constants dans du tampon SDS 1X final (Novex). Le gel est ensuite transféré sur une membrane de nitrocellulose ou de PVDF (tampon de transfert 1X final (Novex) avec 10 % de méthanol) pendant 2 heures à 150 mA constants.

Après transfert, la membrane est bloquée pendant 2 heures à température ambiante
20 dans 50 ml de PBS-T (PBS avec 0.5 % Tween) contenant 2 % de lait écrémé (Merck). L'anticorps primaire (dilué à la concentration optimale de l'ordre du 1/1000e au 1/5000e, dans du PBS-T avec ou sans 2 % de lait écrémé) est laissé sur une nuit à 4 °C. Après un bref rinçage en PBS-T, la membrane est incubée 45 minutes en présence du deuxième anticorps (Ig G anti-souris ou anti-lapin selon le
25 cas, couplé à la peroxydase de raifort) dilué au 1/5000^e dans un tampon dit « ECL » (Tris 20 mM, NaCl 150 mM, Tween 0.1 %).

La membrane est alors rincée 4 fois 15 minutes dans le tampon «ECL». Elle peut être révélée par le réactif ECL (Amersham) constitué de 2 tampons à mélanger extemporanément en volume égal. Différentes expositions d'un film photographique (Hyperfilm ECL ; Amersham) sont effectuées, suivies d'un développement.

5

EXEMPLES

10 **Exemple 1: Isolement de partenaires spécifiques de PS 1 par criblage de banque de fusion d'ADNc de cerveau**

Dans le but d'isoler des partenaires spécifiques de la protéine PS 1, une banque d'ADNc de cerveau humain a été criblée par la technique du double-hybride. Une banque d'ADNc fusionnés avec le DA de GAL4 a été utilisée et différentes protéines appâts ont été construites en fusionnant les régions hydrophiles de PS-I avec le DLA
15 de GAL4.

1.1 Construction des vecteurs permettant l'expression de protéines de fusion entre le DLA de GAL4 et les régions hydrophiles de PS 1.

Cinq protéines de fusion ont été réalisées grâce au vecteur pGBT9 dans lequel ont été introduites les séquences correspondant à différentes régions hydrophiles de PS 1
20 dans le même cadre de lecture que la séquence du DLA de GAL4. Il a été choisi d'identifier des protéines interagissant avec la région N-terminale (Nt, codons 1 à 81), la grande boucle (BL6, codons 263 à 407), la grande boucle avec la mutation L286V (BM), une région plus courte de la grande boucle (BR, codons 290 à 376) et enfin la partie C-terminale de la protéine (Ct, codons 421 à 467).

25 Les fragments d'ADN correspondant à ces différentes régions ont été obtenus par PCR à partir de la séquence entière de l'ADNc de PS 1. Les différents oligonucléotides utilisés ont permis l'introduction des sites EcoRI et SalI aux

extrémités de chaque séquence amplifiée. Ainsi, les fragments de PCR ont pu être insérés entre les sites EcoRI et SalI du multisite de clonage de pGBT9 en aval et en phase avec la séquence du DLA de GAL4. Les différentes constructions obtenues, pGBT-Nt, pGBT-BL6, pGBT-BM, pGBT-BR et pGBT-Ct, ont été vérifiées par séquençage de l'ADN. Cette vérification a montré que les fragments ne présentaient aucune mutation générée par la PCR et qu'ils étaient bien dans le même cadre ouvert de lecture que celui de la séquence du DLA de GAL4.

1.2 Obtention des souches de levures exprimant les protéines de fusion entre le DLA de GAL4 et les régions hydrophiles de PS 1.

- 10 Comme la cotransformation de la levure par deux plasmides différents est un événement moins fréquent que sa transformation par un seul plasmide, une souche de levure préalablement transformée par le vecteur codant pour la protéine appât a été utilisée, afin d'avoir la meilleure efficacité possible lors du criblage de la banque d'ADNc.
- 15 Pour cela, la souche de levure YCM 79 a été transformée par chacun des plasmides pGBT-Nt, pGBT-BL6, pGBT-BM, pGBT-BR et pGBT-Ct. Les levures contenant ces plasmides ont été sélectionnées pour leur phénotype Trp⁺ sur milieu ne contenant pas de tryptophane et respectivement appelées yNt, yBL6, yBM, yBR et yCt. La présence de ces différents plasmides dans chaque souche a été vérifiée en effectuant une PCR
- 20 directement sur les levures, à l'aide d'oligonucléotides complémentaires des séquences DLA et tADH1 situées en 5' et 3' de l'insert. La taille des fragments de PCR obtenus permet de distinguer les différentes souches à l'exception des souches yBL6 et yBM.

1.3 Tests d'activité transactivatrice et d'interaction avec le DA de GAL4 des protéines de fusion contenant les régions hydrophiles de PS 1.

Si une des protéines de fusion présente une activité transactivatrice ou si elle interagit avec le DA de GAL4, elle ne peut pas être utilisée comme appât dans le système

double hybride car elle donne systématiquement une réponse positive en dehors de toute interaction protéine-protéine.

Le premier test consiste, donc, à vérifier que les régions hydrophiles de PS-I ne possèdent pas d'activité transactivatrice lorsqu'elles sont fusionnées au DLA de
5 GAL4, c'est à dire capables d'induire, à elles seules, la transcription des gènes rapporteurs. Ce type d'activité peut, par exemple, être lié à la présence d'un domaine acide dans la protéine de fusion.

Pour ce test, l'expression des gènes rapporteurs *LacZ* et *URA3* ont été contrôlés directement sur les souches yNt, yBL6, yBM, yBR et yCt.

10 Pour le deuxième test qui consiste à contrôler l'absence d'interaction entre les protéines de fusion et le DA de GAL4, les souches yNt, yBL6, yBM, yBR et yCt de phénotype Trp⁺ Leu⁻ ont été transformées par le plasmide pGAD424 (équivalent au pGAD10) qui code pour le DA de GAL4. Les levures ont été sélectionnées sur un milieu sans tryptophane ni leucine et l'expression des gènes rapporteurs *LacZ* et
15 *URA3* permettant de contrôler l'interaction a été testée sur les clones obtenus.

Dans tous les cas, les résultats des tests de l'activité β -galactosidase ainsi que ceux des tests en goutte sur milieu sélectif sans uracile se sont révélés négatifs.

Les protéines de fusions contenant les parties hydrophiles de PS-I ne présentent donc pas de propriété transactivatrice intrinsèque et n'interagissent pas avec le domaine
20 d'activation de GAL4. Elles peuvent donc être utilisées comme protéines appâts pour le criblage de la banque.

1.4 Transformation des différentes souches de levure par la banque de fusion d'ADNc.

Le criblage de la banque de fusion permet d'identifier des clones exprimant des
25 peptides fusionnés au DA de GAL4 interagissant avec la protéine appât. Si cette interaction existe, elle doit permettre la reconstitution d'un transactivateur

fonctionnel capable d'induire l'expression des gènes rapporteurs *URA3* et *LacZ* dans la souche contenant la protéine appât.

Pour que le criblage soit représentatif, il est nécessaire que chaque plasmide indépendant constituant la banque soit présent dans, au moins, un clone de levure après transformation. Pour cela, un protocole permettant théoriquement d'obtenir un nombre de transformants supérieur au nombre de plasmides indépendants de la banque a été utilisé (cf. matériels et méthodes).

Les 5 souches yNt, yBL6, yBM, yBR et yCt de phénotype His-, Lys-, Ade-, Met-, Ura-, Leu-, ont été transformées successivement par 100µg d'ADN plasmidique de la banque de fusion et les cellules transformées ont été sélectionnées sur un milieu YNB +His +Lys +Ade +Met.

A l'issue de la première sélection, nous avons obtenu 32 clones ayant un phénotype Ura+ avec la souche yNt, 450 clones Ura+ avec yBL6, 128 avec yBM, 250 avec yBR et 67 avec yCt. Un test d'activité β -galactosidase a été effectué sur ces transformants afin de vérifier l'expression du deuxième gène rapporteur *LacZ*. Au total 6 clones présentaient le double phénotype Ura+, β -gal+: 1 clone pour la souche yNt, 1 clone pour yBL6, 2 clones pour chacune des souches yBM et yBR, aucun clone pour yCt. Les clones sont respectivement nommés yNt.A, yBL6.B, yBM.C, yBM.D.

EXEMPLE 2 : Isolement des plasmides de la banque.

Afin d'identifier les protéines pouvant interagir avec la boucle mutée ou la boucle réduite de PS 1, les plasmides contenus dans les levures yBM.C et yBM.D ont été extraits. Les plasmides présents dans les préparations d'ADN ont été utilisés pour transformer des bactéries par électroporation.

Les plasmides, extraits des clones bactériens obtenus, ont été digérés par différentes enzymes. Cette analyse a mis en évidence 5 profils de restriction différents.

Les digestions des plasmides extraits des levures yBM.C et yBM.D présentent un profil commun correspondant au plasmide pGBT-BM et deux profils différents correspondant au plasmide pGAD10 contenant un insert de 2900pb correspondant à la séquence ID N° 1 (pGAD-C étant issu de la souche yBM.C) et un insert de 1400pb
5 correspondant à la séquence ID N° 2 (pGAD-D étant issu de la souche yBM.D).

2.1 Contrôle des plasmides isolés par transformation de différentes souches de levure.

Lors de la transformation des levures par les plasmides de la banque de cerveau, il est possible que plusieurs plasmides différents soient introduits dans une même cellule. Afin de vérifier que les résultats d'interaction sont bien dûs à la présence des
10 plasmides isolés et non à des plasmides surnuméraires, nous avons de nouveau testé les protéines de fusion issues du criblage de la banque contre les différentes régions hydrophiles de PS-I dans la souche YCM 79 mais aussi dans les souches HF7 C et PCY 2.

HF7 C possède deux gènes rapporteurs *HIS3* et *LacZ*. Le gène *LacZ* a un
15 environnement promoteur différent de celui existant chez la souche YCM 79, ce qui permet de contrôler que l'obtention de clones β -gal⁺ n'est pas liée à un environnement promoteur spécifique mais bien à une interaction entre les protéines testées. PCY 2 présente l'avantage de ne posséder qu'un seul gène rapporteur *LacZ* qui est exprimé plus fortement que dans les deux autres souches lorsque le
20 transactivateur GAL4 est fonctionnel.

Nous avons transformé les trois souches de levure YCM 79, HF7 C et PCY 2 par les différents couples de plasmides pGBT-X/pGAD-Y (X correspondant à Nt, BL6, BM, BR ou Ct; Y correspondant à C ou D), pGBT9/pGAD424 (témoin d'interaction négatif) et pGBT-CtAPP/pGAD-Fe65 (témoin d'interaction positif) (Russo *et al.*,
25 1995). Nous avons testé l'auxotrophie et l'activité β -gal sur 3 à 5 clones issus de chaque transformation. Les résultats des boîtes sont présentés dans la figure 2.

Compte tenu de la différence de sensibilité des tests, l'interaction est confirmée quand les levures YCM 79 et HF7 C transformées se développent sur les milieux sélectifs sans uracile ou sans histidine, c'est à dire qu'elles expriment, de manière sûre, au moins un de leur gène rapporteur.

- 5 L'intensité de la réponse (colonies plus ou moins bleues, plus ou moins développées) est supposée être proportionnelle au taux d'expression des gènes rapporteurs. Les variations de ce taux peuvent s'expliquer de différentes manières. Un faible taux peut résulter d'une faible interaction entre les protéines reconstituant le transactivateur GAL4. Il peut, aussi, être dû à l'intervention d'une troisième protéine qui
- 10 déstabiliserait la formation du complexe, ou bien à une instabilité des protéines de fusion dans la levure, ou encore à l'état des levures. Les levures peuvent être dans un état de faible activité transcriptionnelle lié, par exemple, à un problème de toxicité des protéines exogènes.

- Les protéines de fusions de la banque de cerveau qui ont été isolées n'interagissent
- 15 pas avec la partie N-terminale et avec le C-terminal de PS 1. Par contre, elles semblent toutes interagir avec la région hydrophile BL6, la région hydrophile BL6 mutée (BM), ainsi que la forme réduite de cette région (BR).

2.2 Tests d'interaction des protéines de fusion de la banque avec des protéines ne contenant pas la boucle de PS 1.

- 20 Ce test permet de déterminer si les protéines de fusion issues de la banque interagissent de façon non spécifique par leur structure, leur charge, leur hydrophobicité avec les protéines appâts (faux positifs de type III) ou avec le DLA de GAL4 (faux positifs de type II).

- Nous avons transformé les souches YCM 79, HF7 C et PCY 2 par les couples de
- 25 plasmides: pGBT9/pGAD-Y (Y correspondant à C ou D), pGBT-LAM/ pGAD-Y, pGBT9/pGAD424 (témoin négatif) et pGBT-CtAPP/pGAD-Fe65 (témoin positif). Le pGBT9 code uniquement pour le DLA de GAL4 alors que le pGBT-LAM fourni par

Clontech code pour une protéine de fusion entre le DLA de GAL4 et la lamine qui interagit de façon non spécifique avec les régions hydrophobes d'autres protéines.

Il en résulte qu'aucune des levures ainsi transformée n'était capable d'exprimer leurs gènes rapporteurs. Ainsi, les levures YCM 79 présentaient le phénotype Ura-, β -gal-,
5 les levures HF7 C étaient toutes His-, β -gal- et les levures PCY 2, β -gal-.

Ces résultats suggèrent que les protéines de fusion de la banque ainsi isolées, interagissent spécifiquement avec la région hydrophile BL6 de PS-I et ne se lient pas avec la lamine ou le DLA de GAL4.

10 2.3 Test d'interaction sur des levures transformées par le vecteur pLex 9 contenant la région codant pour la boucle de PS 1.

Ce test permet de s'assurer que l'interaction des protéines n'est pas due à une structure adoptée par les différents fragments de BL6 uniquement lorsqu'ils sont fusionnés au DLA de GAL4.

Pour vérifier cela, les fragments BL6 et BM ont été fusionnés à un autre DLA, celui
15 du répresseur bactérien LexA. Deux constructions plasmidiques supplémentaires (pLex-BS et pLex-BM) ont été réalisées, où les séquences de la boucle hydrophile BL6 et de la forme mutée ont été fusionnées au gène de *LexA* au niveau des sites EcoRI/SalI du plasmide pLex9.

Ensuite la souche de levure L40 dans laquelle l'expression des gènes rapporteurs
20 *HIS3* et *LacZ* est induite par la fixation du complexe DLA-LexA/DA-GAL4 sur le site opérateur de LexA a été transformée. L'interaction avec les couples de plasmides a été testée : pLex-BL6/pGAD-Y (Y pour C ou D), pLex-BM/pGAD-Y, pLex-RAS/pGAD-RAF (témoin d'interaction positif) (Vojtek *et al.*, 1993) et pLex-RAS/pGAD424, pLex-BL6/pGAD424, pLex-BL6/pGAD-RAF, pLex-
25 BM/pGAD424, pLex-BM/pGAD-RAF (témoins d'interaction négatif).

Les résultats des tests en goutte sur milieux sélectifs sans histidine et des tests de l'activité β -gal se sont révélés positifs pour les souches contenant le plasmide pLex-BL6 quel que soit le plasmide pGAD-Y testé. Il en est de même pour les souches contenant pLex-BM quel que soit le plasmide pGAD-Y testé. Ce test montre que
5 l'interaction obtenue n'est pas liée à une conformation particulière qu'adopterait la région BL6 de PS-I lorsqu'elle est fusionnée avec le domaine de liaison à l'ADN de GAL4.

EXEMPLE 3 : Identification des inserts des plasmides issus du criblage.

3.1 Séquençage des inserts des plasmides pGAD-C et pGAD-D.

- 10 Les inserts des plasmides pGAD-C et pGAD-D ont été entièrement séquencés. Nous avons commencé le séquençage en utilisant les oligonucléotides complémentaires du DA de GAL4 et du tADH1 afin de connaître la séquence du début et de la fin de chaque insert. Puis, à partir de l'extrémité de chaque séquence, nous avons constitué des oligonucléotides de synthèse qui nous ont permis d'avancer dans les séquences
15 sur les deux brins. De cette façon, nous avons séquencé totalement les 1400 pb de l'insert D correspondant à la séquence SEQ ID N°2. Par contre, pour l'insert C de 2900 pb (séquence SEQ ID N°1), il a été nécessaire d'établir une carte de restriction puis d'effectuer un sous clonage de différents fragments afin d'obtenir la totalité de la séquence.
- 20 L'insert D présente une phase ouverte de lecture sur la totalité de sa séquence (Seq ID N°2). En revanche, l'insert C contient un codon stop en fin de séquence (Seq ID N°1).

Leur séquence peptidique présente un profil plutôt hydrophile ce qui est en conformité avec le choix de la méthode de criblage qui s'applique plutôt à des
25 protéines hydrophiles.

3.2 Comparaison des séquences

Cette comparaison doit permettre de savoir si les inserts codent pour un domaine peptidique dont la fonction est connue.

Les séquences nucléiques et peptidiques des inserts ont été comparées avec des séquences répertoriées dans les banques de données « Gen Bank » et « EMBL »
5 (European Molecular Biology Lab). Aucune similitude significative n'a été trouvée entre les inserts et les gènes (ou les protéines) enregistrés dans ces banques.

Les séquences des inserts ont aussi été comparées entre elles. Elles ne présentent aucune similitude.

De plus, l'existence de motifs peptidiques susceptibles d'apporter des informations
10 sur la fonction des peptides codés par les insert a été recherché. Quelques sites potentiels de glycosylation, de phosphorylation et de myristylation ont été trouvés mais qui ne permettent pas de formuler d'hypothèse quant à une éventuelle fonction des peptides.

3.3 Expression des ARNm correspondant aux peptides sélectionnés dans différents 15 tissus humains.

Afin de conforter l'idée que l'interaction entre ces peptides et PS 1 existe bien *in vivo*, il a été vérifié que l'expression des ARNm des peptides isolés de la banque se situe dans les mêmes tissus que l'expression de l'ARNm de PS 1. L'expression dans des
20 tissus spécifiques, des gènes contenant les séquences isolées a été recherchée et le taux d'expression des ARNm correspondant aux insert C et D dans différents tissus humains a été analysé par Northern blot.

La sonde radioactive préparée à partir de l'insert C (séquence ID N°1) a permis d'hybrider un ARNm de 9,5kb exprimé spécifiquement dans le cerveau et plus particulièrement dans le cortex cérébral, le lobe frontal et le lobe temporal (cf. Fig3).
25 La taille de cet ARNm suggère que la protéine traduite est de grande taille.

EXEMPLE 4 : Test d'interaction *in vitro* .

Ce test permet de savoir que les interactions ne sont pas dépendantes du contexte levure. La reconstitution du transactivateur fonctionnel GAL4 n'est pas toujours due à une interaction directe entre les deux protéines, mais peut aussi être le résultat de la formation d'un complexe protéique entre les protéines de fusion et une ou plusieurs protéines endogènes de la levure. Ainsi, les interactions n'existeraient qu'en présence de protéines intermédiaires provenant de la levure et n'auraient pas de réalité en dehors de ce contexte.

Il a été choisi de confirmer les interactions entre la région hydrophile BL6 de PS 1 et les peptides isolés respectivement appelés PepC et PepD dans un système acellulaire.

Les protéines sont produites dans *E.coli* BL21 (DE3) afin d'en obtenir de grande quantité. Les 2 peptides isolés, ont été fusionnés à des peptides tag permettant leur détection et leur fixation sur des supports. Pour cela, il a été choisi de fusionner la protéine GST, qui se lie au glutathion, aux différentes formes de la région BL6 et le peptide S-Tag, reconnu spécifiquement par la protéine S, aux peptides isolés (cf. matériel et méthode pour le principe du test).

4.1 Production d'une protéine de fusion entre le S-Tag et les peptides isolés lors du criblage de la banque d'ADNc.

Afin de construire les protéines de fusion entre le S-Tag et les deux peptides PepC et PepD. Les différents fragments EcoRI/EcoRI issus des plasmides pGAD-C et pGAD-D ont été insérés dans le même cadre de lecture que la séquence de S-Tag du vecteur pET-29a.

A ce stade, un plasmide (pET-D) a été obtenu contenant l'insert correspondant au PepD. Le séquençage partiel de ce plasmide a montré que l'insertion a été réalisée dans la phase de lecture souhaitée.

Ce plasmide a permis l'obtention d'une protéine de 52kDa correspondant à la protéine de fusion STag-PepD. La protéine STag-PepD est produite dans les bactéries BL21 (DE3) après induction à l'IPTG.

4.2 Production d'une protéine de fusion entre la GST et la boucle hydrophile de PS 1.

- 5 Pour fusionner la GST avec l'une des trois formes de la région hydrophile BL6 de PS 1, le vecteur pGEX-4T1 a été utilisé dans lequel les fragments EcoRI/SalI provenant des plasmides pGBT-BL6, pGBT-BM et pGBT-BR ont été introduits.

Le plasmide pGEX-BM correspondant à la boucle mutée de PS-I a été obtenu. Le séquençage a montré que l'insertion a bien été réalisée dans le cadre de lecture de la
10 séquence GST.

Ce plasmide a permis la production de la protéine de fusion GST-BM de 42kDa dans les bactéries BL21 (DE3) après induction à l'IPTG.

4.3 Solubilisation de la protéine de fusion GST-BM.

- Après la lyse des bactéries, la protéine GST-BM reste entièrement dans la fraction
15 insoluble alors que le peptide STag ainsi que la protéine STag-PepD et la GST se trouvent en partie dans la fraction soluble.

La solubilisation des protéines de fusion est absolument nécessaire pour réaliser la chromatographie d'affinité. Dans un premier temps, la protéine GST-BM a été fixée de manière spécifique sur une résine de glutathion couplée de façon covalente à des
20 billes d'agarose afin de pouvoir tester, ensuite, l'interaction de GST-BM avec les partenaires peptidiques de PS 1 fusionné au STag (cf. Matériels et méthodes). Si les protéines ne sont pas solubilisées, elles sédimentent avec les billes en absence de toute liaison spécifique avec celles-ci.

- Après solubilisation de la protéine de fusion GST-BM, le test est réalisé selon les
25 conditions décrites au paragraphe 8 de la section matériels et méthodes.

EXEMPLE 5: Expression des partenaires spécifiques des PS-1 en cellules mammifères.**5 5.1 Construction des vecteurs d'expression appropriés.**

Ce test permet de vérifier si les interactions protéiques révélées en levure sont aussi détectables dans le contexte de cellules mammifères et avec la protéine PS1 complète insérée dans son environnement membranaire. Les séquences SEQ ID N°1 et SEQ ID N°2, des partenaires, ont été sousclonées dans un vecteur d'expression de cellules mammifères pCDNA3 portant une séquence additionnelle correspondant à l'épitope myc en 5' terminale et un site de restriction EcoRI (Fig. 4) permettant le sousclonage en phase des séquences des partenaires.

Le fragment de restriction EcoRI de pGAD-D de 1400 bp a été isolé et souscloné en phase dans le site EcoRI du vecteur d'expression. Les constructions portant l'ADNc dans l'orientation correcte ont été choisies par analyse de restriction enzymatique.

En ce qui concerne le clone C, le fragment EcoRI(1)-HindIII(391) de pGAD-C a d'abord été isolé. Un second fragment HindIII(392)-MscI(2892) (ce dernier site à bout franc se trouvant au sein du vecteur pGAD10 en aval de la région codante du clone C) a ensuite été isolé par digestion partielle. Les fragments EcoRI(1)-HindIII(391) et HindIII(392)-MscI(2892) ont été mis en ligation ensemble avec le vecteur d'expression digéré avec EcoRI et EcoRV, ce dernier site étant à bout franc. Les constructions correctes portant l'intégralité de la séquence codante de C ont été identifiées par analyse de restriction.

Ces constructions permettent de générer des protéines fusion portant l'épitope myc et les peptides correspondant respectivement aux séquences ID N° 1 et ID N° 2.

5.2 Expression dans des cellules de mammifères

Ces constructions ont été transfectées en cellules COS1 par des méthodes standard. Après lyse cellulaire, les échantillons sont analysés par immunotransfert et révélation

avec l'anticorps anti-myc. L'expression des partenaires a pu être détectée par immunotransfert en utilisant l'anticorps anti-Myc 9E10 (fig 5). Nous avons pu détecter une bande très intense de 120 kDa comme attendu pour pepC, démontrant une forte expression (piste C). Pour D, l'intensité de la bande détectée n'était pas
5 toujours constante. Par contre, l'expression en cellules CHO a pu être détectée. Dans ce type cellulaire, D est peut-être un peu plus stable et apparaît comme une protéine migrant vers 80 kDa (piste D).

Les deux protéines ont donc pu être exprimées et détectées en cellules mammifères.

WO 99/16874

PCT/FR98/02093

42

LISTE DE SEQUENCES

5 (1) INFORMATIONS GENERALES:

(i) DEPOSANT:

10 (A) NOM: RHONE-POULENC RORER
 (B) RUE: 20 avenue Raymon Aron
 (C) VILLE: Antony
 (E) PAYS: France
 (F) CODE POSTAL: 92165
 (G) TELEPHONE: 01.55.71.69.22
 (H) TELECOPIE: 01.55.71.72.91

15

(ii) TITRE DE L' INVENTION: Acides nucleiques codant pour des
 proteines capables d'interagir avec les presenilines.

(iii) NOMBRE DE SEQUENCES: 2

20

(iv) FORME DECHIFFRABLE PAR ORDINATEUR:

(A) TYPE DE SUPPORT: Floppy disk
 (B) ORDINATEUR: IBM PC compatible
 (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
 25 (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

(2) INFORMATIONS POUR LA SEQ ID NO: 1:

30 (i) CARACTERISTIQUES DE LA SEQUENCE 1:
 (A) LONGUEUR: 2892 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

35

(ii) TYPE DE MOLECULE: ADNc

(iii) HYPOTHETIQUE: NON

40 (iv) ANTI-SENS: NON

(ix) CARACTERISTIQUE:

(A) NOM/CLE: CDS
 45 (B) EMPLACEMENT: 1..2736

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

50 CGG ATG ATT CCC ATC TTT CAT GAC ATG ATG GAC TGG GAG CAG AGA AAA 48
 Arg Met Ile Pro Ile Phe His Asp Met Met Asp Trp Glu Gln Arg Lys
 1 5 10 15

WO 99/16874

PCT/FR98/02093

43

	AAT GGC AAC TTC AAA CAG GTG GAG GCC GAG TTG ATT GAC AAG CTG GAC	96
	Asn Gly Asn Phe Lys Gln Val Glu Ala Glu Leu Ile Asp Lys Leu Asp	
	20 25 30	
5	AGC ATG GTG TCA GAA GGG AAA GGT GAC GAG AGC TAC AGG GAG CTC TTC	144
	Ser Met Val Ser Glu Gly Lys Gly Asp Glu Ser Tyr Arg Glu Leu Phe	
	35 40 45	
10	AGC CTA CTA ACC CAG CTG TTT GGG CCC TAC CCC AGC CTG CTG GAG AAG	192
	Ser Leu Leu Thr Gln Leu Phe Gly Pro Tyr Pro Ser Leu Leu Glu Lys	
	50 55 60	
15	GTT GAA CAA GAA ACA TGG CGC GAG ACC GGC ATT TCC TTT GTG ACC TCA	240
	Val Glu Gln Glu Thr Trp Arg Glu Thr Gly Ile Ser Phe Val Thr Ser	
	65 70 75 80	
20	GTC ACC CGC CTC ATG GAA CGT CTT CTT GAC TAC AGG GAC TGC ATG AAA	288
	Val Thr Arg Leu Met Glu Arg Leu Leu Asp Tyr Arg Asp Cys Met Lys	
	85 90 95	
25	GGA GAG GAA ACA GAG AAT AAG AAG ATA GGC TGC ACT GTT AAC CTG ATG	336
	Gly Glu Glu Thr Glu Asn Lys Lys Ile Gly Cys Thr Val Asn Leu Met	
	100 105 110	
30	AAT TTT TAC AAA TCT GAG ATT AAC AAG GAA GAA ATG TAT ATC CGC TAC	384
	Asn Phe Tyr Lys Ser Glu Ile Asn Lys Glu Glu Met Tyr Ile Arg Tyr	
	115 120 125	
35	ATC CAT AAG CTT TGT GAC ATG CAC TTG CAG GCC GAA AAC TAC ACA GAG	432
	Ile His Lys Leu Cys Asp Met His Leu Gln Ala Glu Asn Tyr Thr Glu	
	130 135 140	
40	GCC GCA TTT ACC CTG CTC CTT TAC TGT GAG CTG CTG CAG TGG GAG GAC	480
	Ala Ala Phe Thr Leu Leu Leu Tyr Cys Glu Leu Leu Gln Trp Glu Asp	
	145 150 155 160	
45	CGG CCA CTA CGG GAA TTC CTC CAC TAC CCA TCG CAG ACA GAG TGG CAG	528
	Arg Pro Leu Arg Glu Phe Leu His Tyr Pro Ser Gln Thr Glu Trp Gln	
	165 170 175	
50	CGG AAG GAG GGA CTG TGC CGG AAG ATC ATT CAC TAC TTC AAC AAA GGC	576
	Arg Lys Glu Gly Leu Cys Arg Lys Ile Ile His Tyr Phe Asn Lys Gly	
	180 185 190	
55	AAG AGC TGG GAG TTT GGG ATC CCA CTG TGC AGG GAG CTG GCG TGT CAG	624
	Lys Ser Trp Glu Phe Gly Ile Pro Leu Cys Arg Glu Leu Ala Cys Gln	
	195 200 205	
60	TAC GAG AGC CTC TAT GAT TAC CAG AGC CTC AGC TGG ATT CGG AAA ATG	672
	Tyr Glu Ser Leu Tyr Asp Tyr Gln Ser Leu Ser Trp Ile Arg Lys Met	
	210 215 220	

WO 99/16874

PCT/FR98/02093

44

	GAG	GCC	AGC	TAC	TAT	GAC	AAC	ATT	ATG	GAG	CAG	CAA	CGC	CTG	GAG	CCT	720
	Glu	Ala	Ser	Tyr	Tyr	Asp	Asn	Ile	Met	Glu	Gln	Gln	Arg	Leu	Glu	Pro	
	225					230					235					240	
5	GAG	TTC	TTT	CGG	GTC	GGC	TTC	TAT	GGC	AGG	AAG	TTT	CCT	TTC	TTT	CTT	768
	Glu	Phe	Phe	Arg	Val	Gly	Phe	Tyr	Gly	Arg	Lys	Phe	Pro	Phe	Phe	Leu	
					245					250						255	
10	CGG	AAC	AAA	GAA	TAC	GTG	TGC	CGT	GGC	CAT	GAC	TAC	GAG	AGG	CTG	GAG	816
	Arg	Asn	Lys	Glu	Tyr	Val	Cys	Arg	Gly	His	Asp	Tyr	Glu	Arg	Leu	Glu	
				260					265					270			
15	GCC	TTC	CAG	CAG	AGG	ATG	CTC	AGT	GAG	TTT	CCG	CAG	GCT	GTC	GCC	ATG	864
	Ala	Phe	Gln	Gln	Arg	Met	Leu	Ser	Glu	Phe	Pro	Gln	Ala	Val	Ala	Met	
			275					280						285			
20	CAG	CAC	CCC	AAC	CAT	CCT	GAT	GAC	GCC	ATC	CTA	CAG	TGC	GAT	GCC	CAG	912
	Gln	His	Pro	Asn	His	Pro	Asp	Asp	Ala	Ile	Leu	Gln	Cys	Asp	Ala	Gln	
		290					295					300					
25	TAC	TTG	CAG	ATC	TAT	GCA	GTG	ACG	CCC	ATT	CCA	GAT	TAT	GTG	GAT	GTT	960
	Tyr	Leu	Gln	Ile	Tyr	Ala	Val	Thr	Pro	Ile	Pro	Asp	Tyr	Val	Asp	Val	
	305					310					315					320	
30	CTG	CAG	ATG	GAT	AGG	GTA	CCA	GAT	CGA	GTC	AAG	AGC	TTC	TAT	CGC	GTC	1008
	Leu	Gln	Met	Asp	Arg	Val	Pro	Asp	Arg	Val	Lys	Ser	Phe	Tyr	Arg	Val	
					325					330					335		
35	AAC	AAT	GTG	AGG	AAG	TTC	CGG	TAT	GAC	AGG	CCT	TTT	CAC	AAA	GGC	CCC	1056
	Asn	Asn	Val	Arg	Lys	Phe	Arg	Tyr	Asp	Arg	Pro	Phe	His	Lys	Gly	Pro	
				340					345					350			
40	AAG	GAC	AAG	GAG	AAT	GAA	TTC	AAG	AGC	CTG	TGG	ATT	GAA	CGT	ACC	ACA	1104
	Lys	Asp	Lys	Glu	Asn	Glu	Phe	Lys	Ser	Leu	Trp	Ile	Glu	Arg	Thr	Thr	
			355					360					365				
45	CTG	ACC	CTG	ACC	CAC	AGC	TTG	CCT	GGC	ATC	TCT	CGG	TGG	TTT	GAA	GTG	1152
	Leu	Thr	Leu	Thr	His	Ser	Leu	Pro	Gly	Ile	Ser	Arg	Trp	Phe	Glu	Val	
		370					375					380					
50	GAG	AGG	AGG	GAA	CTG	GTG	GAG	GTG	AGC	CCT	CTG	GAG	AAT	GCC	ATC	CAA	1200
	Glu	Arg	Arg	Glu	Leu	Val	Glu	Val	Ser	Pro	Leu	Glu	Asn	Ala	Ile	Gln	
	385					390					395					400	
55	GTG	GTT	GAG	AAT	AAG	AAC	CAG	GAG	CTA	CGC	TCC	CTG	ATC	AGC	CAG	TAT	1248
	Val	Val	Glu	Asn	Lys	Asn	Gln	Glu	Leu	Arg	Ser	Leu	Ile	Ser	Gln	Tyr	
					405					410					415		
60	CAA	CAC	AAG	CAG	GTG	CAT	GGC	AAC	ATT	AAC	CTG	CTA	AGC	ATG	TGC	CTG	1296
	Gln	His	Lys	Gln	Val	His	Gly	Asn	Ile	Asn	Leu	Leu	Ser	Met	Cys	Leu	
				420					425					430			
	AAT	GGT	GTC	ATT	GAT	GCA	GCT	GTC	AAT	GGA	GGC	ATT	GCA	CGC	TAT	CAG	344

WO 99/16874

PCT/FR98/02093

45

	Asn	Gly	Val	Ile	Asp	Ala	Ala	Val	Asn	Gly	Gly	Ile	Ala	Arg	Tyr	Gln	
			435					440					445				
5	GAG	GCC	TTC	TTT	GAT	AAA	GAT	TAC	ATC	AAC	AAG	CAC	CCA	GGA	GAT	GCT	1392
	Glu	Ala	Phe	Phe	Asp	Lys	Asp	Tyr	Ile	Asn	Lys	His	Pro	Gly	Asp	Ala	
		450				455					460						
10	GAG	AAG	ATC	ACC	CAG	CTC	AAG	GAG	CTT	ATG	CAG	GAG	CAG	GTT	CAT	GTC	1440
	Glu	Lys	Ile	Thr	Gln	Leu	Lys	Glu	Leu	Met	Gln	Glu	Gln	Val	His	Val	
	465					470					475					480	
15	CTT	GGA	GTT	GGG	CTA	GCA	GTT	CAT	GAG	AAG	TTT	GTG	CAC	CCA	GAA	ATG	488
	Leu	Gly	Val	Gly	Leu	Ala	Val	His	Glu	Lys	Phe	Val	His	Pro	Glu	Met	
				485						490					495		
	CGG	CCT	CTG	CAT	AAG	AAG	CTA	ATT	GAT	CAG	TTC	CAG	ATG	ATG	CGG	GCC	1536
	Arg	Pro	Leu	His	Lys	Lys	Leu	Ile	Asp	Gln	Phe	Gln	Met	Met	Arg	Ala	
				500					505					510			
20	AGT	CTC	TAC	CAT	GAG	TTT	CCA	GGT	TTG	GAT	AAG	CTA	AGT	CCT	GCA	TGT	1584
	Ser	Leu	Tyr	His	Glu	Phe	Pro	Gly	Leu	Asp	Lys	Leu	Ser	Pro	Ala	Cys	
			515					520					525				
25	TCA	GGC	ACC	AGC	ACC	CCA	CGG	GGA	AAT	GTT	CTG	GCA	TCC	CAT	AGC	CCC	1632
	Ser	Gly	Thr	Ser	Thr	Pro	Arg	Gly	Asn	Val	Leu	Ala	Ser	His	Ser	Pro	
		530					535					540					
30	ATG	AGT	CCG	GAG	AGC	ATC	AAG	ATG	ACC	CAC	CGG	CAC	AGC	CCC	ATG	AAC	1680
	Met	Ser	Pro	Glu	Ser	Ile	Lys	Met	Thr	His	Arg	His	Ser	Pro	Met	Asn	
	545					550					555				560		
35	TTG	ATG	GGC	ACA	GGC	CGC	CAT	TCA	TCA	TCC	TCT	CTC	TCC	TCA	CAT	GCG	1728
	Leu	Met	Gly	Thr	Gly	Arg	His	Ser	Ser	Ser	Ser	Leu	Ser	Ser	His	Ala	
					565					570					575		
	TCT	AGT	GAA	GCA	GGA	AAC	ATG	GTG	ATG	CTG	GGT	GAC	GGC	TCC	ATG	GGT	1776
	Ser	Ser	Glu	Ala	Gly	Asn	Met	Val	Met	Leu	Gly	Asp	Gly	Ser	Met	Gly	
				580				585						590			
40	GAT	GCT	CCT	GAG	GAC	CTG	TAC	CAC	CAC	ATG	CAG	CTC	GCG	TAT	CCC	AAC	1824
	Asp	Ala	Pro	Glu	Asp	Leu	Tyr	His	His	Met	Gln	Leu	Ala	Tyr	Pro	Asn	
			595				600						605				
45	CCC	AGG	TAC	CAA	GGC	TCA	GTC	ACC	AAC	GTC	TCT	GTT	CTG	TCC	TCG	TCC	1872
	Pro	Arg	Tyr	Gln	Gly	Ser	Val	Thr	Asn	Val	Ser	Val	Leu	Ser	Ser	Ser	
		610					615					620					
50	CAG	GCA	AGC	CCT	TCT	TCC	TCC	AGC	CTG	AGT	TCC	ACT	CAC	TCA	GCA	CCA	1920
	Gln	Ala	Ser	Pro	Ser	Ser	Ser	Ser	Leu	Ser	Ser	Thr	His	Ser	Ala	Pro	
	625					630					635					640	
	TCC	CAG	ATG	ATT	ACC	TCT	GCC	CCT	TCC	AGT	GCC	CGA	GAT	AAG	TAC	CGC	1968
	Ser	Gln	Met	Ile	Thr	Ser	Ala	Pro	Ser	Ser	Ala	Arg	Asp	Lys	Tyr	Arg	

WO 99/16874

PCT/FR98/02093

46

	645								650				655				
5	CAT	GCC	CGT	GAA	ATG	ATG	TTG	TTG	CTG	CCC	ACA	TAC	CGG	GAC	CGC	CCA	2016
	His	Ala	Arg	Glu	Met	Met	Leu	Leu	Leu	Pro	Thr	Tyr	Arg	Asp	Arg	Pro	
				660					665					670			
10	AGC	AGT	GCC	ATG	TAT	CCA	GCA	GCC	ATC	CTG	GAG	AAC	GGA	CAG	CCG	CCG	2064
	Ser	Ser	Ala	Met	Tyr	Pro	Ala	Ala	Ile	Leu	Glu	Asn	Gly	Gln	Pro	Pro	
			675					680					685				
	AAT	TTC	CAG	CGA	GCC	CTG	TTC	CAG	CAA	GTG	GTC	GGA	GCC	TGC	AAA	CCC	2112
	Asn	Phe	Gln	Arg	Ala	Leu	Phe	Gln	Gln	Val	Val	Gly	Ala	Cys	Lys	Pro	
		690					695					700					
15	TGC	AGT	GAT	CCC	AAT	CTG	TCT	GTG	GCT	GAA	AAA	GGT	CAT	TAC	TCC	CTA	2160
	Cys	Ser	Asp	Pro	Asn	Leu	Ser	Val	Ala	Glu	Lys	Gly	His	Tyr	Ser	Leu	
	705					710					715					720	
20	CAC	TTT	GAC	GCC	TTC	CAC	CAC	CCT	CTG	GGT	GAT	ACC	CCC	CCA	GCC	CTC	2208
	His	Phe	Asp	Ala	Phe	His	His	Pro	Leu	Gly	Asp	Thr	Pro	Pro	Ala	Leu	
					725					730					735		
25	CCT	GCC	CGG	ACC	CTG	CGC	AAG	TCT	CCT	CTC	CAC	CCT	ATC	CCA	GCC	TCC	2256
	Pro	Ala	Arg	Thr	Leu	Arg	Lys	Ser	Pro	Leu	His	Pro	Ile	Pro	Ala	Ser	
				740					745					750			
30	CCC	ACA	AGC	CCC	CAG	TCA	GGT	CTG	GAC	GGC	AGC	AAC	TCT	ACG	CTG	TCC	2304
	Pro	Thr	Ser	Pro	Gln	Ser	Gly	Leu	Asp	Gly	Ser	Asn	Ser	Thr	Leu	Ser	
			755					760					765				
	GGC	AGT	GCC	AGC	AGC	GGC	GTG	TCC	TCC	TTG	AGT	GAG	AGT	AAC	TTT	GGG	2352
	Gly	Ser	Ala	Ser	Ser	Gly	Val	Ser	Ser	Leu	Ser	Glu	Ser	Asn	Phe	Gly	
		770					775					780					
35	CAC	TCC	TCG	GAG	GCC	CCA	CCT	CGC	ACT	GAC	ACC	ATG	GAC	TCC	ATG	CCA	2400
	His	Ser	Ser	Glu	Ala	Pro	Pro	Arg	Thr	Asp	Thr	Met	Asp	Ser	Met	Pro	
	785					790					795					800	
40	AGT	CAG	GCC	TGG	AAT	GCT	GAC	GAA	GAT	CTT	GAG	CCA	CCC	TAC	CTC	CCT	2448
	Ser	Gln	Ala	Trp	Asn	Ala	Asp	Glu	Asp	Leu	Glu	Pro	Pro	Tyr	Leu	Pro	
					805					810					815		
45	GTC	CAC	TAC	AGC	CTC	TCT	GAG	TCT	GCC	GTC	CTG	GAC	TCC	ATC	AAG	GCC	2496
	Val	His	Tyr	Ser	Leu	Ser	Glu	Ser	Ala	Val	Leu	Asp	Ser	Ile	Lys	Ala	
				820					825					830			
50	CAG	CCA	TGC	CGA	AGC	CAC	TCA	GCC	CCA	GGG	TGC	GTC	ATC	CCT	CAG	GAC	2544
	Gln	Pro	Cys	Arg	Ser	His	Ser	Ala	Pro	Gly	Cys	Val	Ile	Pro	Gln	Asp	
			835					840					845				
	CCC	ATG	GAC	CCG	CCT	GCG	CTG	CCG	CCC	AAG	CCC	TAC	CAC	CCC	CGC	CTG	2592
	Pro	Met	Asp	Pro	Pro	Ala	Leu	Pro	Pro	Lys	Pro	Tyr	His	Pro	Arg	Leu	
		850					855					860					

WO 99/16874

PCT/FR98/02093

47

5 CCG GCC CTG GAG CAC GAT GAG GGG GTG CTG CTG CGT GAA GAG ACT GAG 2640
 Pro Ala Leu Glu His Asp Glu Gly Val Leu Leu Arg Glu Glu Thr Glu
 865 870 875 880

AGG CCT CGA GGC CTG CAC CGC AAG GCT CCA TTG CCT CCT GGG AGC GCT 2688
 Arg Pro Arg Gly Leu His Arg Lys Ala Pro Leu Pro Pro Gly Ser Ala
 885 890 895

10 AAG GAG GAG CAG GCC CGC ATG GCC TGG GAG CAC GGC CGA GGG GAG CAG 2736
 Lys Glu Glu Gln Ala Arg Met Ala Trp Glu His Gly Arg Gly Glu Gln
 900 905 910

15 TGAGGGGCAA CGAGGCGGCT GGGATGCCGC CCTCAGTAAG CAGCTTGCCA ATCACTCCAG
 2796

GTCTGAAAAG CAAGTCCCCC AGCCCCACCC CAGGGAGCCA GAGAGGCTTG CACTCAGGAG
 2856

20 AGAACCACCC CCAAGTCTCC GTTCTACTGC CGCCCG 2892

WO 99/16874

PCT/FR98/02093

48

(2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: 1363 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

10 (ii) TYPE DE MOLECULE: ADNc

(iii) HYPOTHETIQUE: NON

15 (iv) ANTI-SENS: NON

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
 (B) EMBLACEMENT: 1..1363
 20

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

25	GAA TTC CGG GGT TGT GTG AAG ATG GAG TTT CCC GGA GGA AAT GAC AAT	48
	Glu Phe Arg Gly Cys Val Lys Met Glu Phe Pro Gly Gly Asn Asp Asn	
	915 920 925	
30	TAC CTG ACG ATC ACA GGG CCT TCG CAC CCC TTC CTG TCA GGG GCC GAG	96
	Tyr Leu Thr Ile Thr Gly Pro Ser His Pro Phe Leu Ser Gly Ala Glu	
	930 935 940	
35	ACA TTC CAT ACA CCA AGC TTG GGT GAT GAG GAA TTT GAA ATC CCA CCT	144
	Thr Phe His Thr Pro Ser Leu Gly Asp Glu Glu Phe Glu Ile Pro Pro	
	945 950 955 960	
40	ATC TCC TTG GAT TCT GAT CCC TCA TTG GCT GTC TCA GAT GTG GTT GGC	192
	Ile Ser Leu Asp Ser Asp Pro Ser Leu Ala Val Ser Asp Val Val Gly	
	965 970 975	
45	CAC TTT GAT GAC CTG GCA GAC CCT TCC TCT TCA CAG GAT GGC AGT TTT	240
	His Phe Asp Asp Leu Ala Asp Pro Ser Ser Ser Gln Asp Gly Ser Phe	
	980 985 990	
50	TCA GCC CAG TAT GGG GTC CAG ACA TTG GAC ATG CCT GTG GGC ATG ACC	288
	Ser Ala Gln Tyr Gly Val Gln Thr Leu Asp Met Pro Val Gly Met Thr	
	995 1000 1005	
	CAT GGC TTG ATG GAG CAG GGC GGG GGG CTC CTG AGT GGG GGC TTG ACC	336
	His Gly Leu Met Glu Gln Gly Gly Gly Leu Leu Ser Gly Gly Leu Thr	
	1010 1015 1020	
	ATG GAC TTG GAC CAC TCT ATA GGA ACT CAG TAT AGT GCC AAC CCA CCT	384
	Met Asp Leu Asp His Ser Ile Gly Thr Gln Tyr Ser Ala Asn Pro Pro	

WO 99/16874

PCT/FR98/02093

49

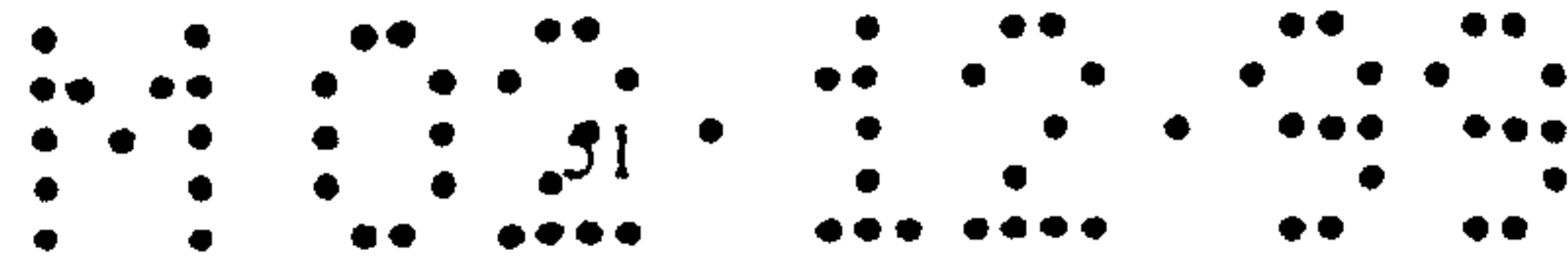
	1025	1030	1035	1040	
5	GTT ACA ATT GAT GTA CCA ATG ACA GAC ATG ACA TCT GGC TTG ATG GGG Val Thr Ile Asp Val Pro Met Thr Asp Met Thr Ser Gly Leu Met Gly	1045	1050	1055	432
10	CAT AGC CAG TTG ACC ACC ATT GAT CAG TCA GAA CTG AGT TCC CAG CTG His Ser Gln Leu Thr Thr Ile Asp Gln Ser Glu Leu Ser Ser Gln Leu	1060	1065	1070	480
15	GGT TTG AGC CTA GGG GGT GGC ACC ATC CTG CCA CCT GCC CAG TCA CCT Gly Leu Ser Leu Gly Gly Gly Thr Ile Leu Pro Pro Ala Gln Ser Pro	1075	1080	1085	528
20	GAA GAT CGT CTT TCA ACC ACC CCT TCA CCT ACT AGT TCA CTT CAC GAA Glu Asp Arg Leu Ser Thr Thr Pro Ser Pro Thr Ser Ser Leu His Glu	1090	1095	1100	576
25	GAT GGT GTT GAG GAT TTC CGG AGG CAA CTT CCC AGC CAG AAG ACA GTC Asp Gly Val Glu Asp Phe Arg Arg Gln Leu Pro Ser Gln Lys Thr Val	1105	1110	1115	624
30	GTG GTG GAA GCA GGG AAA AAG CAG AAG GCC CCA AAG AAG AGA AAA AAG Val Val Glu Ala Gly Lys Lys Gln Lys Ala Pro Lys Lys Arg Lys Lys	1125	1130	1135	672
35	AAA GAT CCT AAT GAA CCT CAG AAA CCA GTT TCA GCA TAT GCT TTA TTC Lys Asp Pro Asn Glu Pro Gln Lys Pro Val Ser Ala Tyr Ala Leu Phe	1140	1145	1150	720
40	TTT CGT GAT ACA CAG GCT GCC ATC AAG GGA CAG AAT CCT AAT GCC ACT Phe Arg Asp Thr Gln Ala Ala Ile Lys Gly Gln Asn Pro Asn Ala Thr	1155	1160	1165	768
45	TTT GGT GAG GTT TCA AAA ATT GTG GCC TCC ATG TGG GAT AGT CTT GGA Phe Gly Glu Val Ser Lys Ile Val Ala Ser Met Trp Asp Ser Leu Gly	1170	1175	1180	816
50	GAG GAG CAA AAA CAG GTA TAT AAG AGG AAA ACT GAG GCT GCC AAG AAA Glu Glu Gln Lys Gln Val Tyr Lys Arg Lys Thr Glu Ala Ala Lys Lys	1185	1190	1195	864
	GAG TAT CTG AAG GCA CTG GCT GCT TAC AAA GAC AAC CAG GAG TGT CAG Glu Tyr Leu Lys Ala Leu Ala Ala Tyr Lys Asp Asn Gln Glu Cys Gln	1205	1210	1215	912
	GCC ACT GTG GAA ACA GTG GAA TTG GAT CCA GCA CCA CCA TCA CAA ACT Ala Thr Val Glu Thr Val Glu Leu Asp Pro Ala Pro Pro Ser Gln Thr	1220	1225	1230	960
	CCT TCT CCA CCT CCT ATG GCT ACT GTT GAC CCA GCA TCT CCA GCA CCA Pro Ser Pro Pro Pro Met Ala Thr Val Asp Pro Ala Ser Pro Ala Pro	1235	1240	1245	1008

WO 99/16874

PCT/FR98/02093

50

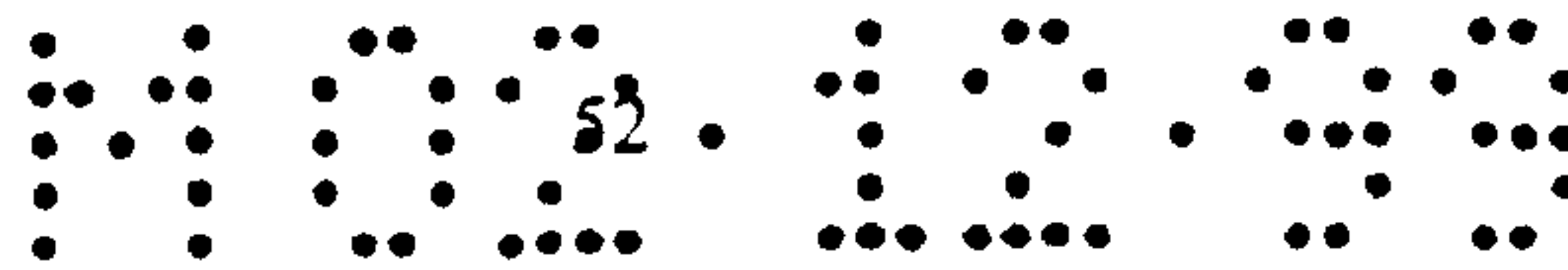
	GCT TCA ATA GAG CCC CCT GCC CTG TCC CCA TCC ATT GTT GTT AAC TCC	1056
	Ala Ser Ile Glu Pro Pro Ala Leu Ser Pro Ser Ile Val Val Asn Ser	
	1250 1255 1260	
5	ACC CTT TCA TCC TAT GTG GCA AAC CAG GCA TCT TCT GGA GCT GGG GGT	1104
	Thr Leu Ser Ser Tyr Val Ala Asn Gln Ala Ser Ser Gly Ala Gly Gly	
	1265 1270 1275 1280	
10	CAG CCC AAT ATC ACC AAG TTG ATT ATT ACC AAA CAA ATG TTG CCC TCT	1152
	Gln Pro Asn Ile Thr Lys Leu Ile Ile Thr Lys Gln Met Leu Pro Ser	
	1285 1290 1295	
15	TCT ATT ACT ATG TCT CAA GGA GGG ATG GTT ACT GTT ATC CCA GCC ACA	1200
	Ser Ile Thr Met Ser Gln Gly Gly Met Val Thr Val Ile Pro Ala Thr	
	1300 1305 1310	
20	GTG GTG ACC TCC CGG GGG CTC CAA CTA GGC CAA ACC AGT ACA GCT ACT	1248
	Val Val Thr Ser Arg Gly Leu Gln Leu Gly Gln Thr Ser Thr Ala Thr	
	1315 1320 1325	
25	ATC CAG CCC AGT CAA CAA GCC CAG ATT GTC ACT CGG TCA GTG TTG CAG	1296
	Ile Gln Pro Ser Gln Gln Ala Gln Ile Val Thr Arg Ser Val Leu Gln	
	1330 1335 1340	
	GCA GCA GCA GCT GCT GCT GCT GCT GCT TCT ATG CAA CTG CCT CCA CCC	1344
	Ala Ala Ala Ala Ala Ala Ala Ala Ala Ser Met Gln Leu Pro Pro Pro	
	1345 1350 1355 1360	
30	CGA GTA CAG CCC GGA ATT C	1363
	Arg Val Gln Pro Gly Ile	
	1365	



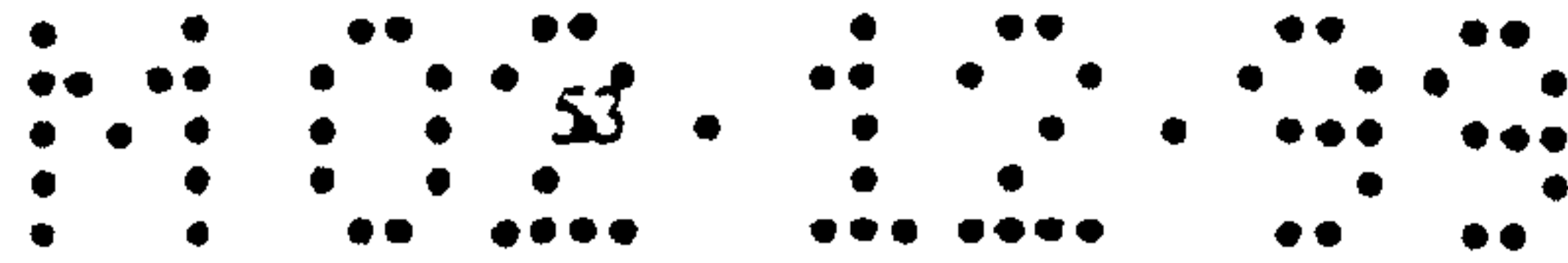
REVENDICATIONS

1. Séquence nucléotidique comprenant tout ou partie de la séquence telle que définie dans la séquence SEQ ID N°1 et codant pour une protéine caractérisée en ce qu'elle est capable d'interagir avec les présénilines.
2. Séquence nucléotidique comprenant tout ou partie de la séquence SEQ ID N°2 ou de ses dérivées et codant pour une protéine caractérisée en ce qu'elle est capable d'interagir avec les présénilines.
3. Polypeptides comprenant tout ou partie d'une séquence peptidique choisie parmi la séquence SEQ ID N°1 telle que définie dans la revendication 1 ou la séquence SEQ ID N°2 ou de ses dérivés définie dans la revendication 2 et caractérisés en ce qu'ils sont capables d'interagir avec les présénilines.
4. Polypeptides selon la revendication 3 caractérisés en ce qu'il sont capables d'interagir avec la grande boucle de la préséniline PS-1.
5. Polypeptides selon la revendication 4 caractérisés en ce qu'ils sont capables d'interagir avec la séquence d'acides-amino 310-463 de la grande boucle de la préséniline PS-1.
6. Polypeptides selon la revendication 3 caractérisés en ce qu'ils sont capables d'interagir avec la grande boucle de la préséniline PS-2.
7. Utilisation d'un polypeptide selon les revendications 3 à 6 pour l'obtention d'un médicament destiné à ralentir, stimuler ou moduler l'activité des présénilines.

FEUILLE MODIFIEE



8. Procédé de préparation d'un polypeptide selon l'une des revendications 3 à 6 caractérisé en ce que l'on cultive une cellule contenant une séquence nucléotidique selon l'une des revendications 1 ou 2 dans des conditions d'expression de ladite séquence et on récupère le polypeptide produit
9. Cellule hôte pour la production de polypeptide selon l'une des revendications 3 à 6, caractérisée en ce qu'elle a été transformée avec un acide nucléique comportant une séquence nucléotidique selon l'une des revendications 1 ou 2.
10. Oligonucléotide antisens d'une séquence selon l'une des revendications 1 ou 2.
11. Sonde nucléotidique capable de s'hybrider avec une séquence nucléotidique selon l'une des revendications 1 ou 2 ou l'ARNm correspondant
12. Anticorps ou fragment d'anticorps dirigé contre un polypeptide selon l'une des revendications 3 à 6.
13. Anticorps ou fragment d'anticorps selon la revendication 12 caractérisé en ce qu'il est dirigé contre une séquence choisie parmi les séquences peptidiques présentées en SEQ ID N° 1 ou SEQ ID N°2 ou leurs dérivées.
14. Anticorps ou fragment d'anticorps selon la revendication 12 ou 13 caractérisé en ce qu'il possède la faculté d'inhiber et/ ou de révéler l'interaction entre les présénilines et un polypeptide selon l'une des revendications 3 à 6.
15. Procédé de mise en évidence ou d'isolement de composés capables de moduler ou d'inhiber l'interaction entre les présénilines et les polypeptides tels que définis dans les revendication 3 à 6, caractérisé en ce que l'on réalise les étapes suivantes :
 - a - on met en contact une molécule ou un mélange contenant différentes



molécules, éventuellement non-identifiées, avec une cellule recombinée telle que décrite ci-dessus exprimant un polypeptide de l'invention dans des conditions permettant l'interaction entre ledit polypeptide et ladite molécule dans le cas où celle-ci posséderait une affinité pour ledit polypeptide, et,

b - on détecte et/ou isole les molécules liées au dit polypeptide.

16. Ligand d'un polypeptide tel que défini selon les revendications 3 à 6, susceptible d'être obtenu selon le procédé de la revendication 15.

17. Utilisation d'un ligand selon la revendication 16 pour la préparation d'un médicament destiné au traitement des affections neurologiques.

18. Virus recombinant defectif comprenant une séquence nucléotidique codant pour un polypeptide selon l'une des revendications 3 à 6.

19. Vecteur comprenant une séquence nucléotidique selon l'une des revendications 1 ou 2 .

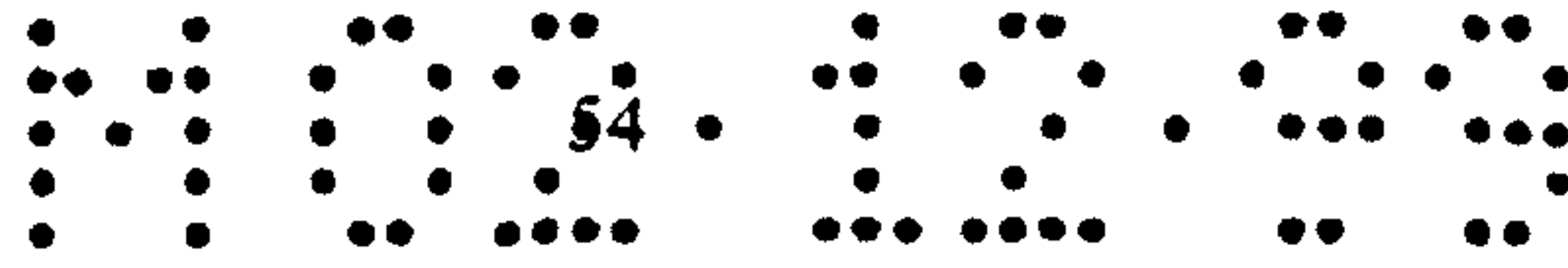
20. Vecteur selon la revendication 19 caractérisé en ce qu'il s'agit d'un vecteur plasmidique.

21. Vecteur selon la revendication 20 caractérisé en ce qu'il s'agit d'un vecteur viral.

22. Vecteur selon la revendication 20 caractérisé en ce qu'il s'agit d'un virus defectif pour la réplication.

23. Composition pharmaceutique comprenant un ou plusieurs vecteurs selon l'une des revendications 18 à 22.

FEUILLE MODIFIEE



24. Composition pharmaceutique comprenant comme principe actif au moins un polypeptide selon l'une des revendications 3 à 6.
25. Composition pharmaceutique comprenant comme principe actif au moins un anticorps ou fragment d'anticorps selon l'une des revendications 12 à 14, et/ou un oligonucléotide antisens selon la revendication 10 et/ou un composé préparé selon la revendication 15.
26. Composition selon l'une des revendications 23 à 25 destinée à moduler ralentir ou inhiber l'activité des présénilines ou de leurs formes mutées.
27. Composition selon l'une des revendications 23 à 25 destinée au traitement des maladies neurodégénératives.

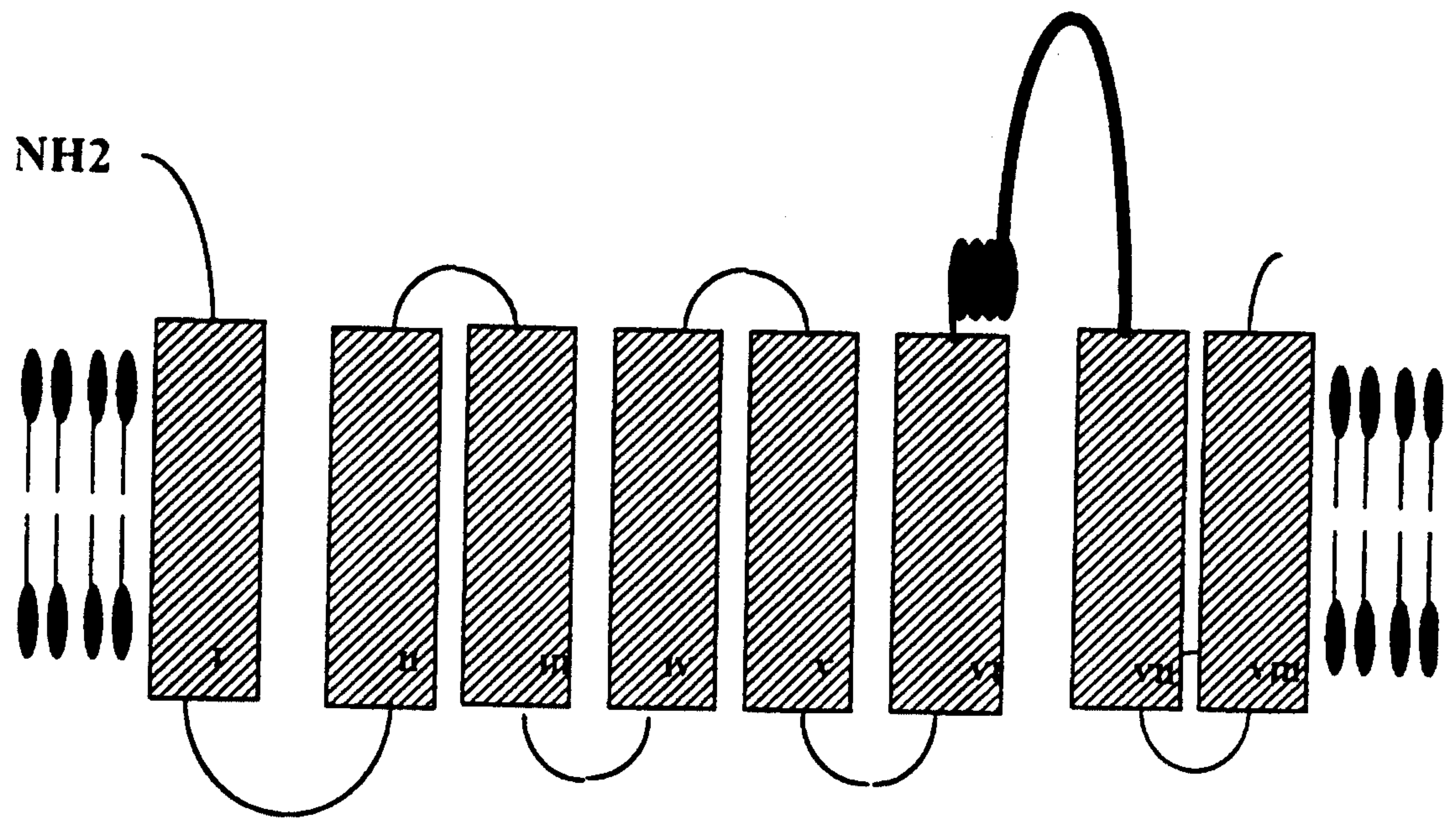
CYTOPLASME**LUMIERE DU RETICULUM
ENDOPLASMIQUE**

Figure 1. Structure schématique des présénilines

L'appât utilisé pour PS1 est indiqué en gras

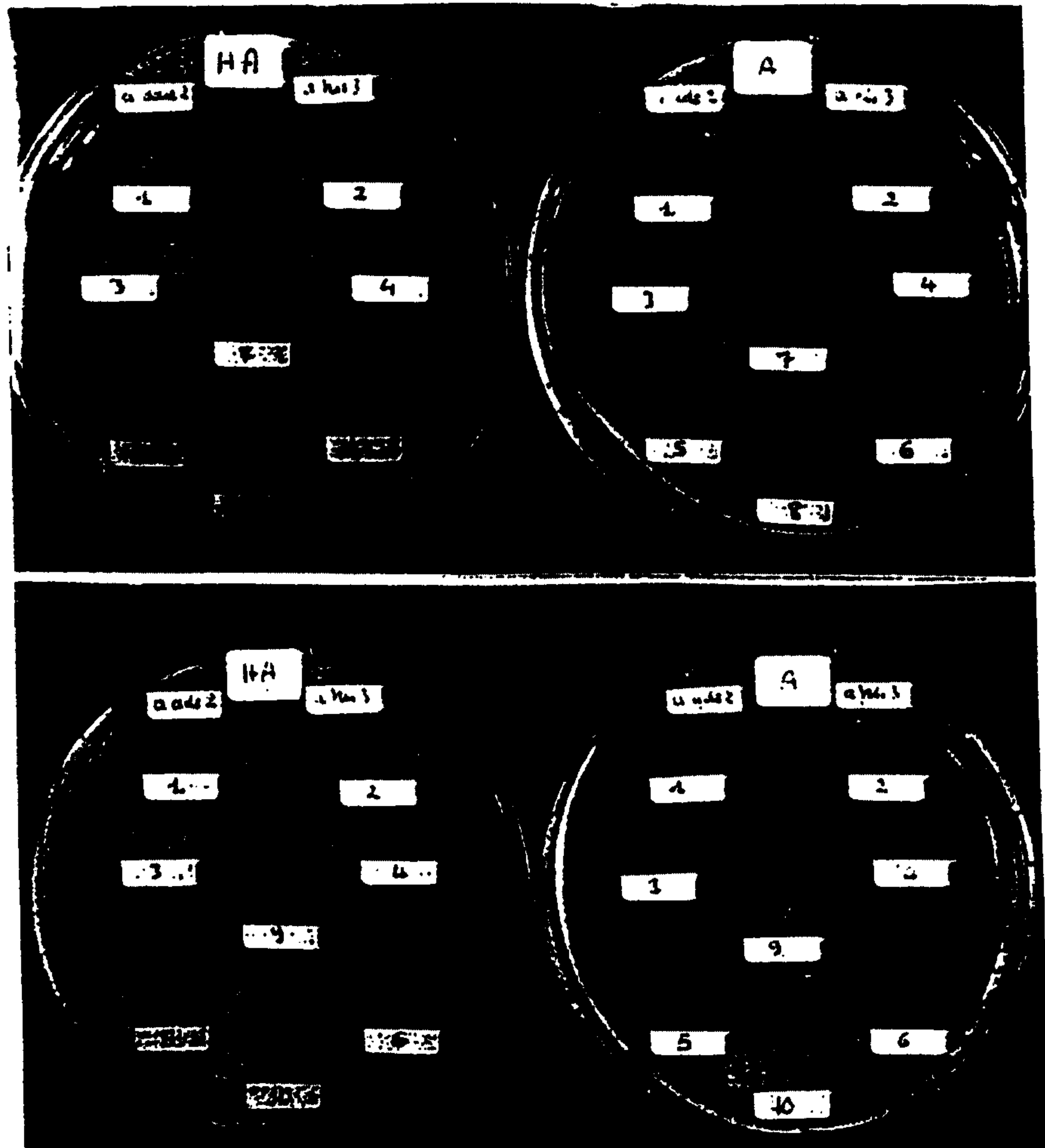


Figure.2 : Test des souches yBM-C et yBM-D

- 1 Cerebellum

2 Cortex cérébral

3 Medulla

4 Pôle occipital

5 Lobe frontal

6 Lobe temporal

7 Putamen

8 Corde spinale
- 9 Coeur

10 Cerveau

11 Placenta

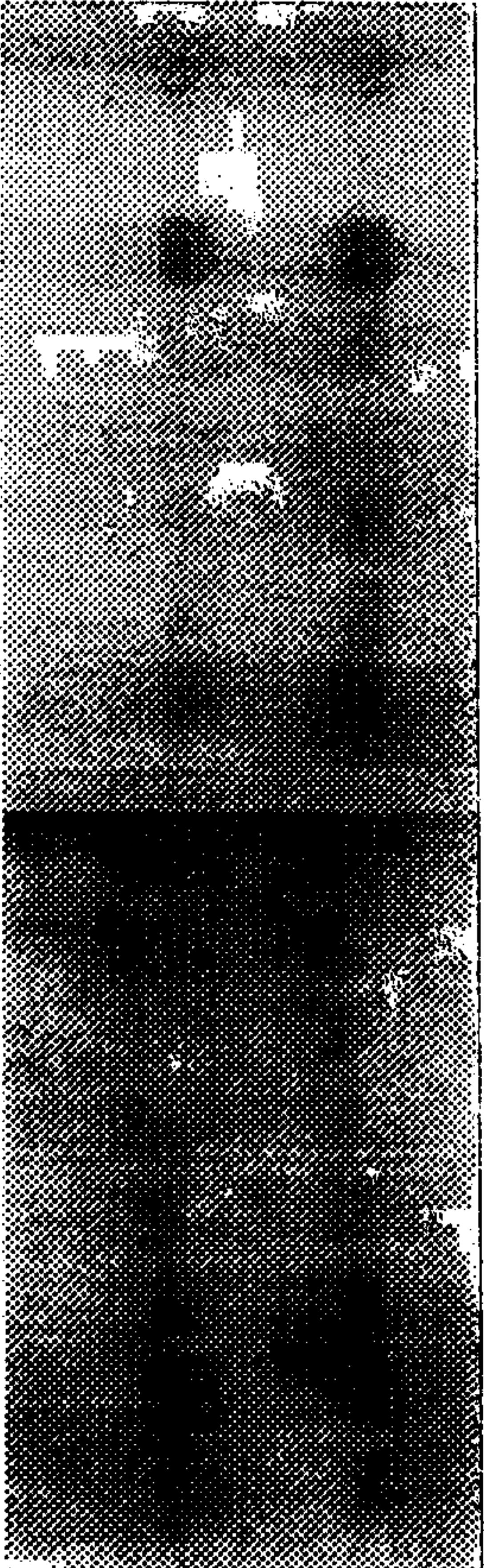
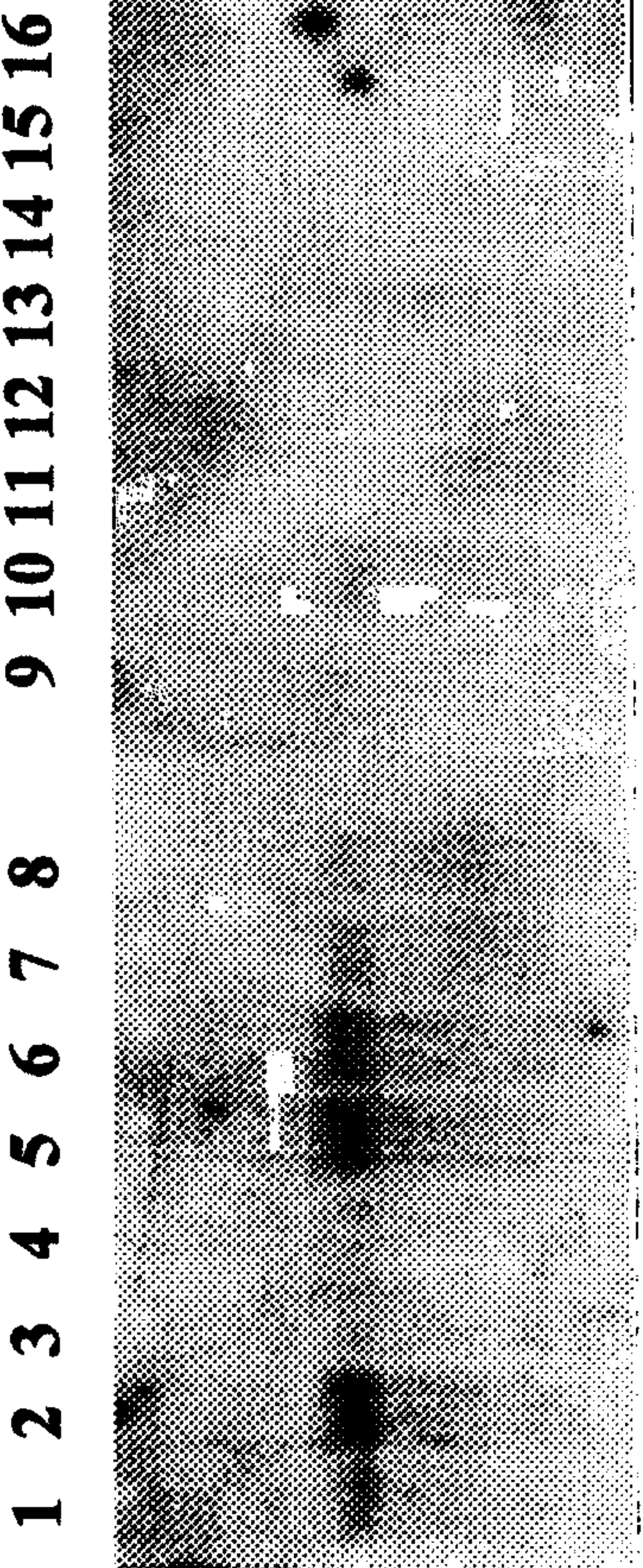
12 Poumon

13 Foie

14 Muscle squelettique

15 Reins

16 Pancréas



SONDE C

SONDE D

Figure 3

FEUILLE DE REMPLACEMENT (REGLE 26)

4/8

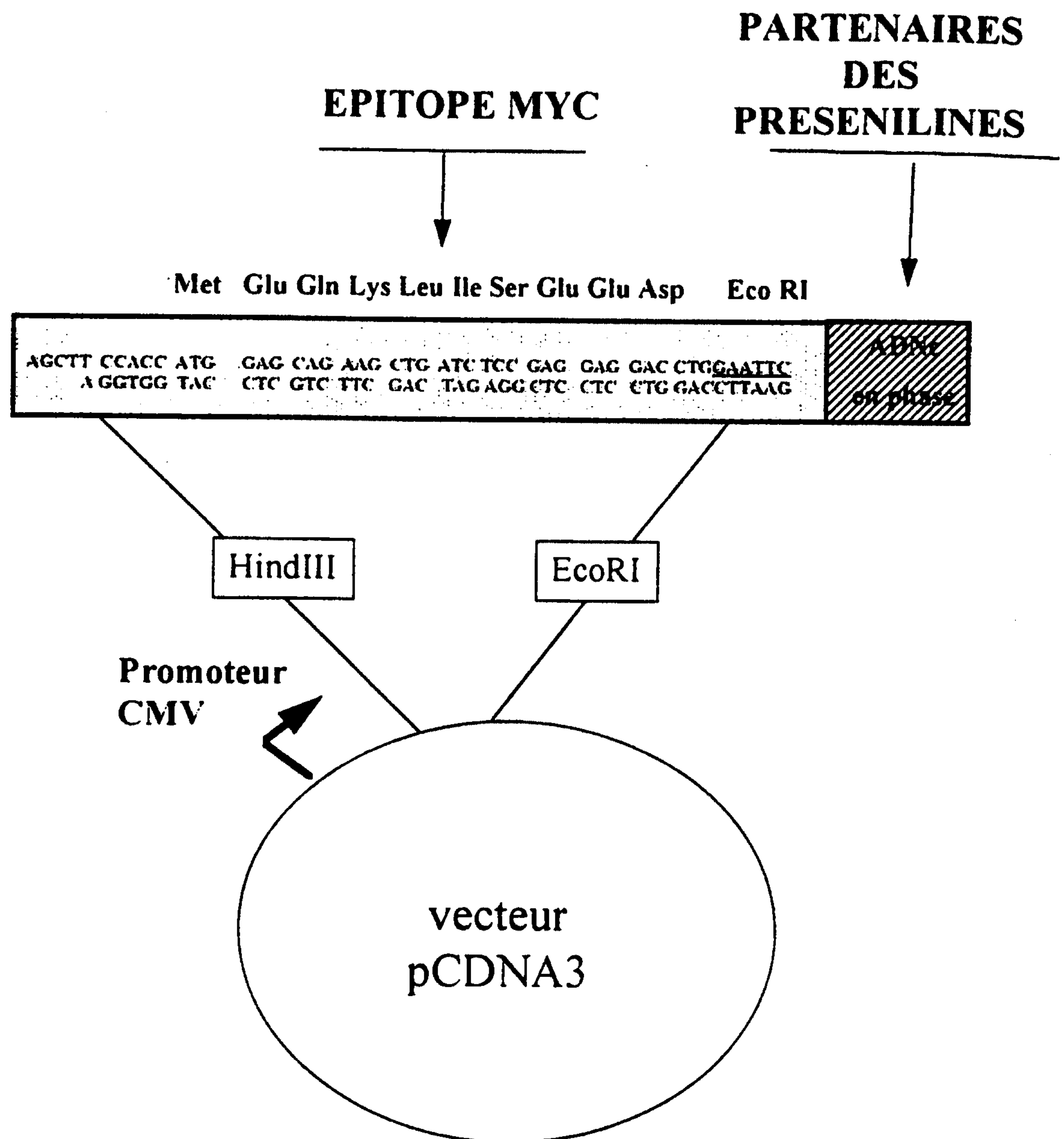


Figure 4. Schéma du vecteur d'expression
mammifère utilisé.

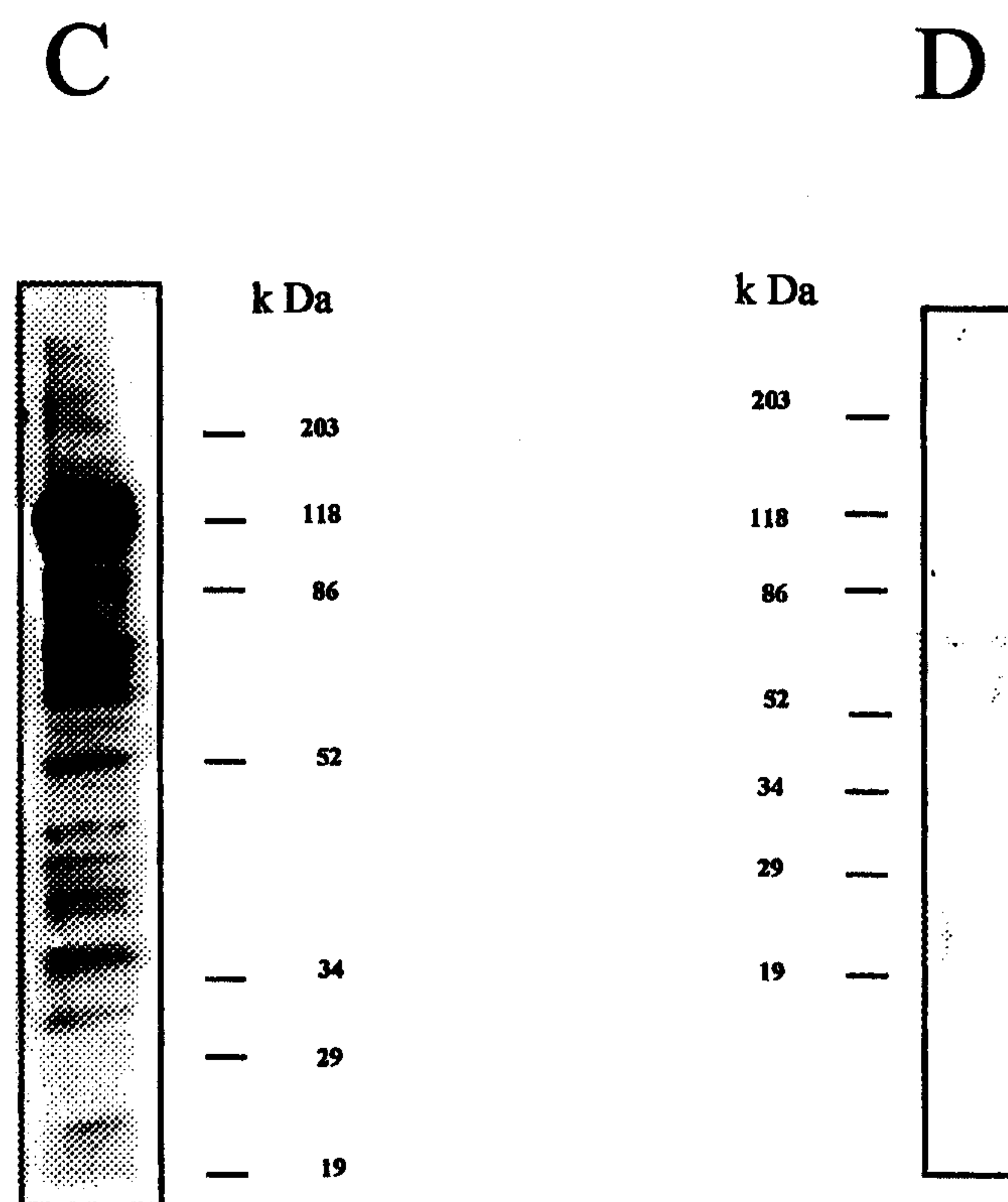
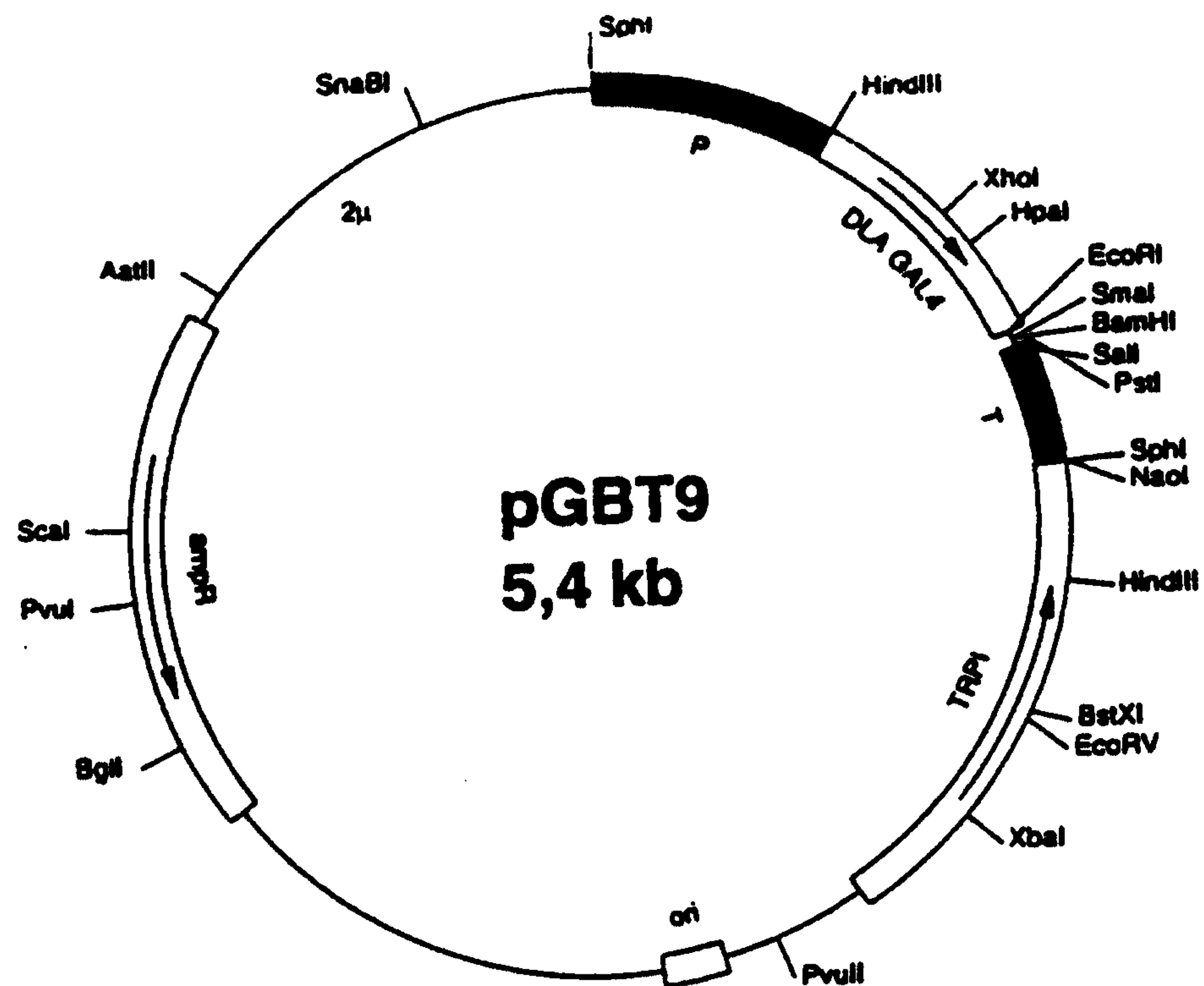
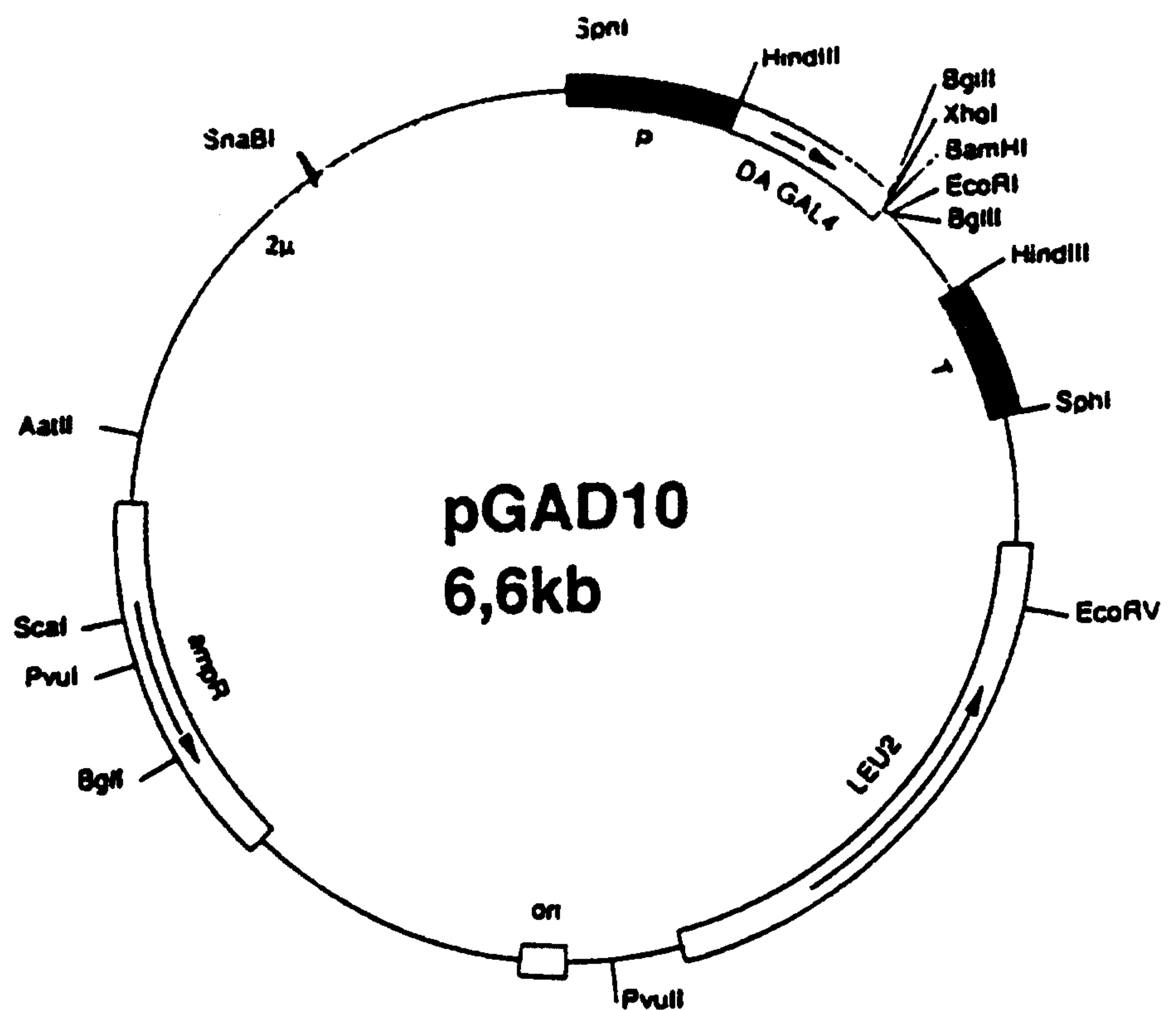


Figure 5. Expression des protéines fusion C et D en cellules mammifères

5



10



15

Figure 6 (a): Cartes des principaux vecteurs utilisés
FEUILLE DE REMPLACEMENT (REGLE 26)

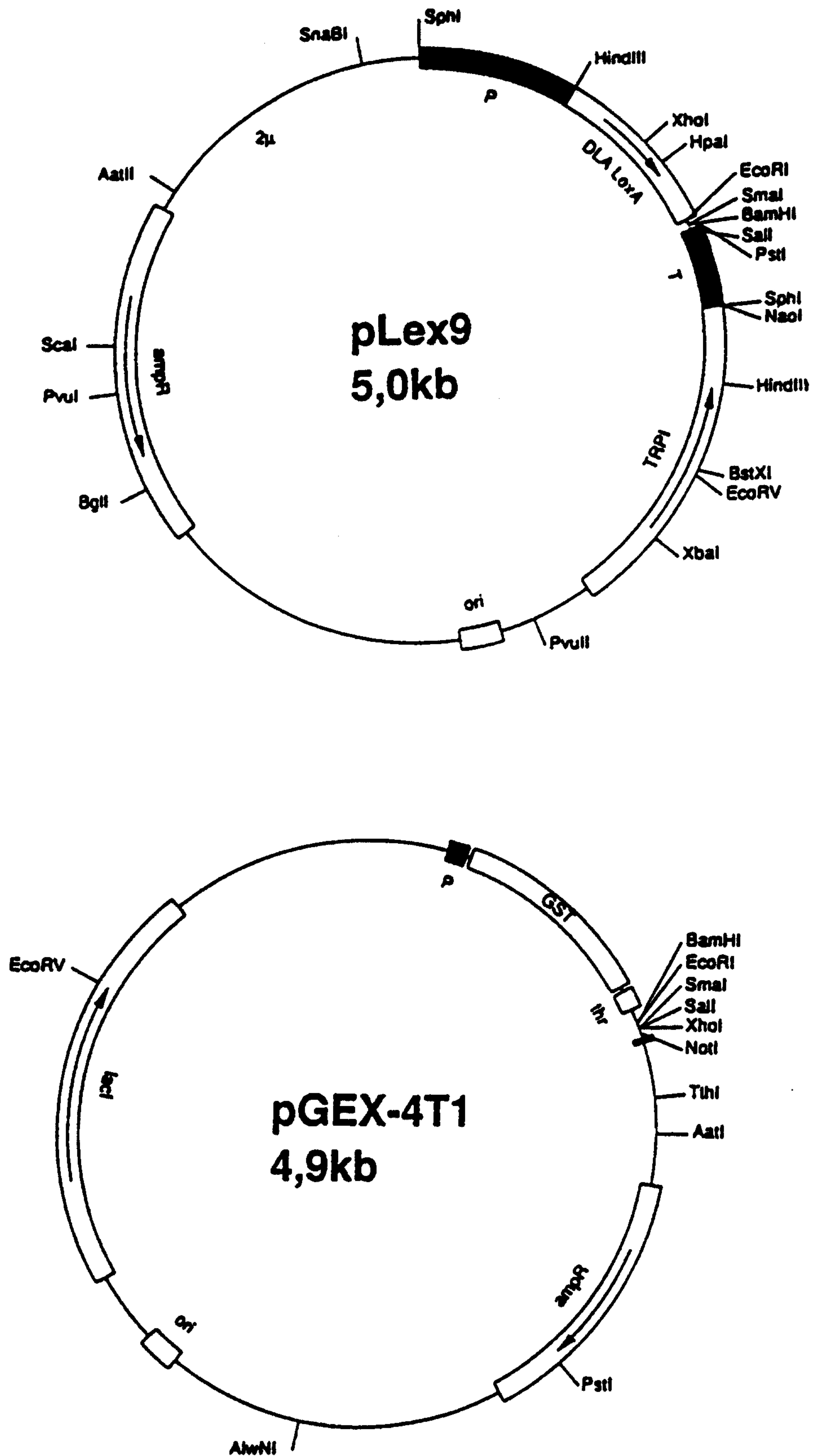


Figure 6 (b): Cartes des principaux vecteurs utilisés

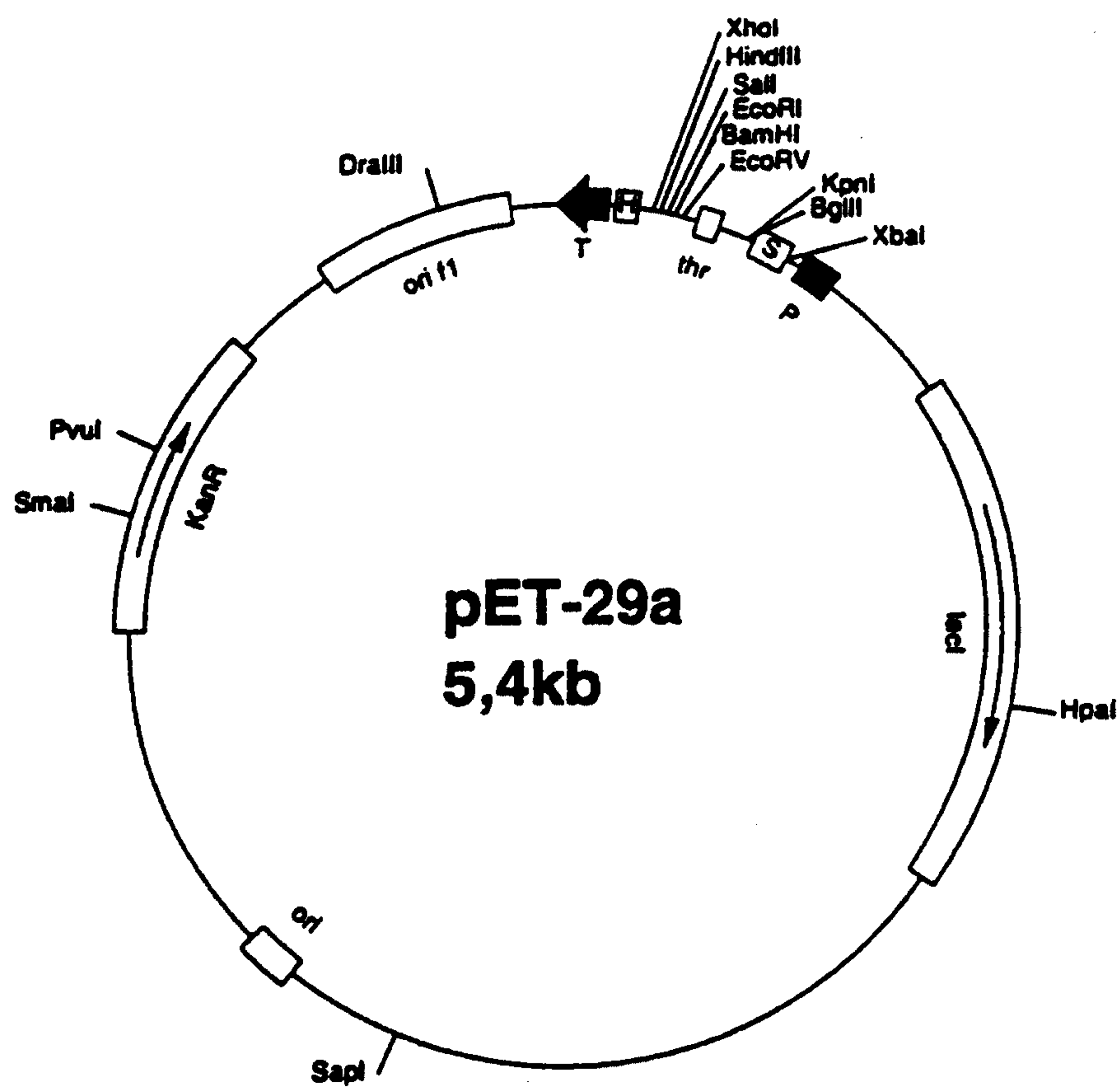


Figure 6 (c): Cartes des principaux vecteurs utilisés